SUMEX

STANFORD UNIVERSITY
MEDICAL EXPERIMENTAL COMPUTER RESOURCE

RR-00785

ANNUAL REPORT—YEAR 13

Submitted to

BIOMEDICAL RESEARCH TECHNOLOGY PROGRAM
NATIONAL INSTITUTES OF HEALTH

June 1, 1986

STANFORD UNIVERSITY SCHOOL OF MEDICINE
Edward H. Shortliffe, Principal Investigator
Edward A. Feigenbaum, Co-Principal Investigator

DEPARTMENT OF HEALTH AND HUMAN SERVICES
PUBLIC HEALTH SERVICE
NATIONAL INSTITUTES OF HEALTH

DIVISION OF RESEARCH RESOURCES
BIOMEDICAL RESEARCH TECHNOLOGY PROGRAM

ANNUAL PROGRESS REPORT
PART I., TITLE PAGE

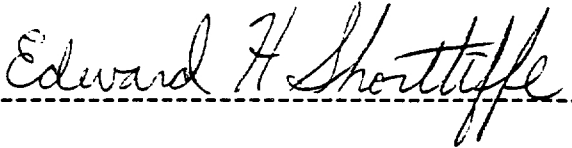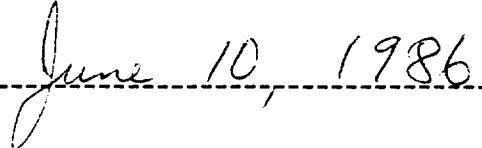1. PHS GRANT NUMBER: 5 P41 RR00785-13

2. TITLE OF GRANT: SUMEX

                                               Stanford University Medical
                                               Experimental Computer Resource

3. NAME OF RECIPIENT INSTITUTION: Stanford University

4. HEALTH PROFESSIONAL SCHOOL: School of Medicine

5. REPORTING PERIOD:

    5a.   FROM: 08-01-85
    5b.   TO: 07-31-86

6. PRINCIPAL INVESTIGATOR:

    6a.   NAME: Edward H. Shortliffe, M.D., Ph.D.
    6b.   TITLE: Associate Professor of Medicine
                                               and Computer Science

    6c.   SIGNATURE: *Edward H Shortliffe*

7. DATE SIGNED: *June 10, 1986*

8. TELEPHONE: 415-723-6979

Table of Contents

E. H. Shortliffe

List of Figures

E. H. Shortliffe

II. Description of Program Activities

This section corresponds to the predefined forms required by the Division of Research Resources to provide information about our resource activities for their computerized retrieval system. These forms have been submitted separately and are not reproduced here to avoid redundancy with the more extensive narrative information about our resource and progress provided in this report.

II.A. Scientific Subprojects

Our core research and development activities are described starting on page 14, our training activities are summarized starting on page 53, and the progress of our collaborating projects is detailed starting on page 93.

II.B. Books, Papers, and Abstracts

The list of recent publications for our core research and development work starts on page 45 and those for the collaborating projects are in the individual reports starting on page 93.

II.C. Resource Summary Table

The details of resource usage, including a breakdown by the various subprojects, is given in the tables starting on page 56.

III. Narrative Description

III.A. Summary of Research Progress

III.A.1. Overview

It is now thirteen years since the SUMEX-AIM resource was established in 1973 and before discussing the details of our progress this past year, we take this opportunity to reflect on the broad progress of SUMEX as a resource. Computing and communications technologies and biomedical artificial intelligence (AI) research have achieved remarkable results. The SUMEX-AIM resource has both profoundly influenced and responded to those changing technologies. It is widely recognized that our resource has fostered highly influential work in biomedical AI -- work from which much of the field of expert systems has emerged -- and that it has simultaneously helped define the technological base of applied AI research.

SUMEX has been the home of such well-known AI systems as DENDRAL (chemical structure elucidation), MYCIN (infectious disease diagnosis and therapy), INTERNIST (differential diagnosis), ACT (human memory organization), ONCOCIN (cancer chemotherapy protocol advice), SECS (chemical synthesis), EMYCIN (rule-based expert system tool), and AGE (blackboard-based expert system tool). In the past four years, our community has published a dozen books that give a scholarly perspective on the scientific experiments we have been performing. These volumes, and other work done at SUMEX, have played a seminal role in structuring modern AI paradigms and methodology.

SUMEX has the reputation of a model national resource, pulling together the best available interactive computing technology, software, and computer communications in the service of a national scientific community. Planning groups for national facilities in cognitive science, computer science, and biomathematical modeling have discussed and studied the SUMEX model and new resources, like the recently instituted BIONET resource for molecular biologists, are closely patterned after the SUMEX example. SUMEX has demonstrated that a computer resource is a useful "linking mechanism" for bringing together and holding together teams of experts from different disciplines who share a common problem focus. For example, computer scientists have been collaborating fruitfully with physical chemists, molecular biochemists, geneticists, crystallographers, internists, ophthalmologists, infectious disease specialists, intensive care specialists, oncologists, psychologists, biomedical engineers, and other expert practitioners. And in some of these cases, the interdisciplinary collaboration, usually so difficult to achieve in the best of circumstances, was achieved in spite of geographical distance between the participants, using the computer networks.

SUMEX has also achieved successes as a community builder. AI concepts and software are among the most complex products of computer science. Historically it has not been easy for scientists in other fields to gain access to and mastery of them. Yet the collaborative outreach and dissemination efforts of SUMEX have been able to bridge the gap in numerous cases. Over 36 biomedical AI application projects have developed in our national community and have been supported by SUMEX over the years. And 9 of these have matured to the point of now continuing their research on facilities outside of SUMEX. For example, the BIONET resource (named GENET while at SUMEX) is being operated by IntelliCorp; the Rutgers Computers in Biomedicine resource is centered at Rutgers University; the CADUCEUS project splits their research work between their own VAX computer and the SUMEX resource; and the Chemical

Synthesis project now operates entirely on a VAX at U.C. Santa Cruz. Interest in AI research and application continues to grow. AI is one of the principal fronts along which university computer science groups are expanding. Federal and industrial support for AI research is vigorous and growing, although support specifically for biomedical applications continues to be limited. Nevertheless, there is an explosion of interest in medical AI. The American Association for Artificial Intelligence (AAAI), the principal scientific membership organization for the AI field, has 7000 members, over 1000 of whom are members of the medical special interest group known as the AAAI-M. Speakers on medical AI are prominently featured at professional medical meetings, such as the American College of Pathology and American College of Physicians meetings; a decade ago, the words "artificial intelligence" were never heard at such conferences. And at medical computing meetings, such as the annual Symposium on Computer Applications in Medical Care and the international MEDINFO conferences, the growing interest in AI and the rapid increase in papers on AI and expert systems are further testimony to the impact that the field is having.

AI is beginning to have a similar effect on medical education. Such diverse organizations as the National Library of Medicine, the American College of Physicians, the Association of American Medical Colleges, and the Medical Library Association have all called for sweeping changes in medical education, increased educational use of computing technology, enhanced research in medical computer science, and career development for people working at the interface between medicine and computing. They all cite evolving computing technology and (SUMEX-AIM) AI research as key motivators.

Even as we reflect on this substantial progress, however, at the deepest research level the problems we can attack are still sharply limited. Our current ideas fall short in many ways against today's important health care and biomedical research problems brought on by the explosion in medical knowledge and for which AI should be of assistance. Just as the research work of the 70's and 80's in the SUMEX-AIM community fuels the current practical and commercial applications, our work of the late 80's will be the basis for the next decade's systems. Our growing knowledge is clearly attained in an incremental fashion; we build today on the results of the past decade, and we will build in the 1990's on the work we undertake today.

At the resource level, there is a growing, diverse, and active AIM research community with needs for more and more powerful computing resources to continue its work. Many of these groups still are dependent on the SUMEX-AIM resources. For those who have been able to take advantage of newly developed local computing facilities, SUMEX-AIM provides a central cross-roads for communications and the sharing of programs and knowledge. In its core research and development role, SUMEX-AIM has its sights set on the hardware and software systems of the next decade. We expect major changes in the distributed computing environments that are just now emerging in order to make effective use of their power and to adapt them to the development and dissemination of biomedical AI systems for professional user communities. This has been the major focus of our core system research this past year.

In its training role, SUMEX is a crucial resource for the education of badly needed new researchers and professionals to continue the development of the biomedical AI field. The "critical mass" of the existing physical SUMEX resource, its development staff, and its intellectual ties with the Stanford Knowledge Systems Laboratory, make this an ideal setting to integrate, experiment with, and export these methodologies for the rest of the AIM community.

III.A.2. Resource Goals and Definitions

SUMEX-AIM is a national computer resource with a multiple mission: a) promoting experimental applications of computer science research in artificial intelligence (AI) to biological and medical problems, b) studying methodologies for the dissemination of biomedical AI systems into target user communities, c) supporting the basic AI research that underlies applications, and d) facilitating network-based computer resource sharing, collaboration, and communication among a national scientific community of health research projects. The SUMEX-AIM resource is located physically in the Stanford University Medical School and serves as a nucleus for a community of medical AI projects at universities around the country. SUMEX provides computing facilities tuned to the needs of AI research and communication tools to facilitate remote access, inter- and intra-group contacts, and the demonstration of developing computer programs to biomedical research collaborators.

III.A.2.1. What is Artificial Intelligence?

Artificial Intelligence research is that part of Computer Science concerned with symbol manipulation processes that produce intelligent action [1, 6, 7, 8]. Here *intelligent action* means an act or decision that is goal-oriented, is arrived at by an understandable chain of symbolic analysis and reasoning steps, and utilizes knowledge of the world to inform and guide the reasoning.

Placing AI in Computer Science

A simplified view relates AI research with the rest of computer science. The manner of use of computers by people to accomplish tasks can be thought of as a one-dimensional spectrum representing the nature of the instructions that must be given the computer to do its job. At one extreme of the spectrum, representing early computer science, users supply their intelligence to instruct the machine precisely *how* to do the job, step-by-step.

At the other extreme of the spectrum, users describe *what* they wish the computer to do for them to solve problems. They want to communicate what is to be done without having to lay out in detail all necessary subgoals for adequate performance, yet with a reasonable assurance that they are addressing an intelligent agent that is using knowledge of their world to understand their intent, complain or fill in their vagueness, make specific their abstractions, correct their errors, discover appropriate subgoals, and ultimately translate what they want done into detailed processing steps that define how it should be done by a real computer. Users want to provide this specification of what to do in a language that is comfortable to them and the problem domain (perhaps English) and via communication modes that are convenient (including perhaps speech or pictures).

Progress in computer science may be seen as steps away from that extreme *how* point on the spectrum: the familiar panoply of assembly languages, subroutine libraries, compilers, extensible languages, etc. illustrate this trend. The research activity aimed at creating computer programs that act as *intelligent agents* near the *what* end of the spectrum can be viewed as a long-range goal of AI research.

E. H. Shortliffe

Expert Systems and Applications

The national SUMEX-AIM resource has enabled a long, interdisciplinary line of artificial intelligence research at Stanford concerned with the development of concepts and techniques for building expert systems [3]. An *expert system* is an intelligent computer program that uses knowledge and inference procedures to solve problems that are difficult enough to require significant human expertise for their solution. For some fields of work, the knowledge necessary to perform at such a level, plus the inference procedures used, can be thought of as a model of the expertise of the expert practitioners of that field.

The knowledge of an expert system consists of facts and heuristics. The *facts* constitute a body of information that is widely shared, publicly available, and generally agreed upon by experts in a field. The *heuristics* are the mostly-private, little-discussed rules of good judgment (rules of plausible reasoning, rules of good guessing) that characterize expert-level decision making in the field. The performance level of an expert system is primarily a function of the size and quality of the knowledge base that it possesses.

Projects in the SUMEX-AIM community are concerned in some way with the application of AI to biomedical research. Brief abstracts of the various projects currently using the SUMEX resource can be found in Appendix B and more detailed progress summaries in Section IV. The most tangible objective of this approach is the development of computer programs that will be more general and effective consultative tools for the clinician and medical scientist. There have already been promising results in areas such as chemical structure elucidation and synthesis, diagnostic consultation, molecular biology, and modeling of psychological processes.

Needless to say, much is yet to be learned in the process of fashioning a coherent scientific discipline out of the experimental programs, mathematical procedures, and emerging theoretical structure comprising artificial intelligence research. State-of-the-art programs are far more narrowly specialized and inflexible than the corresponding aspects of human intelligence they emulate; however, in special domains they may be of comparable or greater power, e.g., in the solution of structure problems in organic chemistry or in the rigorous consideration of a large diagnostic knowledge base.

III.A.2.2. Resource Sharing

An equally important function of the SUMEX-AIM resource is an exploration of the use of computer communications as a means for interactions and sharing between geographically remote research groups engaged in biomedical computer science research and for the dissemination of AI technology. This facet of scientific interaction is becoming increasingly important with the explosion of complex information sources and the regional specialization of groups and facilities that might be shared by remote researchers [5, 2]. And, as projected, we are seeing a growing decentralization of computing resources with the emerging technology in microelectronics and a correspondingly greater role for digital communications to facilitate scientific exchange.

Our community building effort is based upon the developing state of distributed computing and communications technology. While far from perfected, these capabilities offer powerful tools for collaborative linkages, both within a given research project and among them. A number of the active projects on SUMEX are based upon the collaboration of computer and medical scientists at geographically separate institutions, separate both from each other and from the computer resource (see for example, the MENTOR and PathFinder projects).

In the early 1970's, the initial model for SUMEX-AIM as a centralized resource was based on the high cost of powerful computing facilities and the infeasibility of being able to duplicate them readily. This central role has already evolved significantly and continues to change with the introduction of more compact and inexpensive computing technology now available at many more research sites. At the same time, the number of active groups working on biomedical AI problems has grown and the established ones have increased in size. This has led to a growth in the demand for computing resources far beyond what SUMEX-AIM could reasonably and effectively provide on a national scale. We have actively supported efforts by the more mature AIM projects to develop or adapt additional computing facilities tailored to their particular needs and designed to free the main SUMEX resource for new, developing applications projects. To date, over 9 of the national projects have moved some or all of their work to local sites and several have begun resource communities of their own (see page 87). Thus, as more remotely available resources have become established, the balance of the use of the SUMEX-AIM resource has shifted toward supporting start-up pilot projects and the growing AI research community at Stanford.

E. H. Shortliffe

III.A.2.3. Significance to Biomedicine

Artificial intelligence is the computer science of representations of symbolic knowledge and its use in symbolic inference and problem-solving processes. There is a certain inevitability to this branch of computer science and its applications, in particular, to medicine and biosciences. The cost of computers will continue to fall drastically during the coming two decades. As it does, many more of the practitioners of the world's professions will be persuaded to turn to economical automatic information processing for assistance in managing the increasing complexity of their daily tasks. They will find, from most of computer science, help only for those problems that have a mathematical or statistical core, or are of a routine data-processing nature. But such problems will be relatively rare, except in engineering and physical science. In medicine, biology, management, indeed in most of the world's work, the daily tasks are those requiring symbolic reasoning with detailed professional knowledge. The computers that will act as *intelligent assistants* for these professionals must be endowed with symbolic reasoning capabilities and knowledge.

The growth in medical knowledge has far surpassed the ability of a single practitioner to master it all, and the computer's superior information processing capacity thereby offers a natural appeal. Furthermore, the reasoning processes of medical experts are poorly understood; attempts to model expert decision-making necessarily require a degree of introspection and a structured experimentation that may, in turn, improve the quality of the physician's own clinical decisions, making them more reproducible and defensible. New insights that result may also allow us more adequately to teach medical students and house staff the techniques for reaching good decisions, rather than merely to offer a collection of facts which they must independently learn to utilize coherently.

The knowledge that must be used is a combination of factual knowledge and heuristic knowledge. The latter is especially hard to obtain and represent since the experts providing it are mostly unaware of the heuristic knowledge they are using. Medical and scientific communities currently face many widely-recognized problems relating to the rapid accumulation of knowledge, for example:

- codifying theoretical and heuristic knowledge

- effectively using the wealth of information implicitly available from textbooks, journal articles and other practitioners

- disseminating that knowledge beyond the intellectual centers where it is collected

- customizing the presentation of that knowledge to individual practitioners as well as customizing the application of the information to individual cases

We believe that computers are an inevitable technology for helping to overcome these problems. While recognizing the value of mathematical modeling, statistical classification, decision theory and other techniques, we believe that effective use of such methods depends on using them in conjunction with less formal knowledge, including contextual and strategic knowledge.

Artificial intelligence offers advantages for representing and using information that will allow physicians and scientists to use computers as intelligent assistants. In this way we envision a significant extension to the decision-making powers of specific practitioners without reducing the importance of those individuals in that process.

Knowledge is power, in the profession and in the intelligent agent. As we proceed to model expertise in medicine and its related sciences, we find that the power of our

programs derives mainly from the knowledge that we are able to obtain from our collaborating practitioners, not from the sophistication of the inference processes we observe them using. Crucially, the knowledge that gives power is not merely the knowledge of the textbook, the lecture and the journal, but the knowledge of *good practice* -- the experiential knowledge of *good judgment* and *good guessing*, the knowledge of the practitioner's art that is often used in lieu of facts and rigor. This heuristic knowledge is mostly private, even in the very public practice of science. It is almost never taught explicitly, is almost never discussed and critiqued among peers, and most often is not even in the moment-by-moment awareness of the practitioner.

Perhaps the the most expansive view of the significance of the work of the SUMEX-AIM community is that a methodology is emerging for the systematic explication, testing, dissemination, and teaching of the heuristic knowledge of medical practice and scientific performance. Perhaps it is less important that computer programs can be organized to use this knowledge than that the knowledge itself can be organized for the use of the human practitioners of today and tomorrow.

In summary, the logic which mandates that artificial intelligence play a key role in enhancing knowledge management and access for biomedicine -- a logic in which we have long believed -- has gradually become evident to much of the biomedical community. We are encouraged by this increased recognition, but humbled by the realization of the significant research challenges that remain. Our goals are accordingly both scientific and educational. We continue to pursue the research objectives that have always guided SUMEX-AIM, but must also undertake educational efforts designed to inform the biomedical community of our results while cautioning it about the challenges remaining.

E. H. Shortliffe

III.A.2.4. Summary of Current Goals

The following summarizes SUMEX-AIM resource objectives as stated in the proposal for the on-going five-year grant, begun on August 1, 1981, and provides the backdrop against which specific progress is reported. These project goals are presented in the three categories used in the previous proposal: 1) resource operations, 2) training and education, and 3) core research.

1) Resource Operations

- Maintain the vitality of the AIM community by continuing to encourage and explore new applications of AI to biomedical research and improving mechanisms for inter- and intra-group collaborations and communications. User projects will fund their own manpower and local needs; will actively contribute their special expertise to the SUMEX-AIM community; and will receive an allocation of computing resources under the control of the AIM management committees. There will be no "fee for service" charges for community members.

- Provide effective computational support for AIM community goals, including efforts to improve the support for artificial intelligence research and new applications work; to develop new computational tools to support more mature projects; and to facilitate testing and research dissemination of nearly operational programs. We will continue to operate and develop the existing. mainframe facility as the nucleus of the resource. We will acquire additional equipment to meet developing community needs for more capacity, larger program address spaces, and improved interactive facilities. New computing hardware technologies becoming available now and in the next few years will play a key role in these developments and we expect to take the lead in this community for adapting these new tools to biomedical AI needs.

- Provide effective and geographically accessible communication facilities to the SUMEX-AIM community for remote collaborations, communications among distributed computing nodes, and experimental testing of AI programs. We will retain the current ARPANET and TYMNET connections for at least the near term and will actively explore other advantageous connections to new communications networks and to dedicated links.

2) Training and Education

- Provide community-wide support and work to make resource goals and AI programs known and available to appropriate medical scientists. Collaborating projects are responsible for the development and dissemination of their own AI programs.

- Provide documentation and assistance to interface users to resource facilities and programs and continue to exploit particular areas of expertise within the community for developing pilot efforts in new application areas.

- Allocate "collaborative linkage" funds to qualifying new and pilot projects to provide for communications and terminal support pending formal approval and funding of their projects. These funds are allocated in cooperation with the AIM Executive Committee reviews of prospective user projects.

- Support workshop activities, including collaboration with the Rutgers

Computers in Biomedicine resource on the AIM community workshop and with individual projects for more specialized workshops covering specific application areas or program dissemination.

3) Core Research

- Explore basic artificial intelligence research issues and techniques, including knowledge acquisition, representation, and utilization; reasoning in the presence of uncertainty; strategy planning; and explanations of reasoning pathways, with particular emphasis on biomedical applications.

- Support community efforts to organize and generalize AI tools that have been developed in the context of individual application projects. This will include work to organize the present state-of-the-art in AI techniques through the AI Handbook effort and the development of practical software packages (e.g., AGE, EMYCIN, UNITS, and EXPERT) for the acquisition, representation, and utilization of knowledge in AI programs.

E. H. Shortliffe

III.A.3. Details of Technical Progress

This progress summary covers the nucleus of the SUMEX-AIM resource. Objectives and progress for individual collaborating projects are discussed in their respective reports in Section IV. These collaborative projects collectively provide much of the scientific basis for SUMEX as a resource and our role in assisting them has been a continuation of that evolved in the past. Collaborating projects are autonomous in their management and provide their own manpower and expertise for the development and dissemination of their AI programs.

III.A.3.1. Progress Highlights

In this section we summarize highlights of SUMEX-AIM resource activities over the past year (May 1985 - April 1986), focusing on the resource nucleus.

- We have made additional significant improvements to the SUMEX-AIM computing environment in order to optimize computing support for the community. These include the addition of 18 Xerox 1186 workstations, 20 Texas Instruments Explorer workstations, and 4 Symbolics workstations. The purchase of these Lisp machines was funded jointly by NIH, DARPA, and machine vendor gifts. We continue to operate the mainframe computers (DEC 2060, 2020, VAX 11/780, and VAX 11/750's) for the community. Because of the broad mix of research in the SUMEX-AIM community, no single computer vendor can meet our needs so we have undertaken long-term support of a heterogeneous computing environment, incorporating many types of machines linked through multiprotocol Ethernet facilities.

- We have continued the core development of the SUMEX software tools and networking systems to enhance the facilities available to researchers. Much of this work has centered on the effective integration of distributed computing resources in the form of mainframes, workstations, and servers. Network gateways and terminal interface machines have been enhanced with new protocol and internet routing capabilities. We have developed many new software packages to enhance the computing environments of the Lisp workstations and to link them to other hosts and servers on our networks.

- We have continued the dissemination of SUMEX-AIM technology through various media. We have distributed 73 copies of our AI software tools (EMYCIN, AGE, MRS, SACON, and BB1) to academic, industrial, and federal research laboratories. We have also continued to distribute the video tapes of some of our research projects including ONCOCIN, and an overview tape of Knowledge Systems Laboratory work to outside groups. The KSL overview tape won a CINE Eagle award for excellence and a recommendation to represent U.S. scientific efforts at events abroad.

- Our group has continued to publish actively on the results of our research, including more than 45 research papers per year in the AI literature and a dozen books in the past 5 years on various aspects of SUMEX-AIM AI research (see page 89).

- The Medical Information Sciences program, begun at Stanford in 1983 under Professor Shortliffe as Director, has continued to grow over the past year to include about 25 outstanding PhD and MS students, and a search for an additional faculty member has been authorized and is underway. The specialized curriculum offered by the MIS program focuses on the

development of a new generation of researchers able to support the development of improved computer-based solutions to biomedical needs. The feasibility of this program resulted in large part from the prior work and research computing environment provided by the SUMEX-AIM resource. Over 20 PhD and MS trainees will be enrolled in the fall of 1985. It has been awarded post-doctoral training support from the National Library of Medicine, received equipment gifts from Xerox and Hewlett-Packard, and has received additional industrial and foundation grants for student support.

- We made significant progress in core AI research. In the area of knowledge representation, further work was done on the representation of explicit strategy knowledge, temporal knowledge, causal knowledge, and knowledge in logic-based systems. In the area of architectures and control, we worked on a new implementation of a blackboard architecture with explicit control knowledge. Under knowledge acquisition studies, work continues on experiments in learning by induction, by analogy, and learning from partial theories. In the area of knowledge utilization, results include work on reasoning with uncertainty and using counterfactual conditionals. We continued work on a number of existing tools for expert systems and on building new ones such as the BB1 system.

- We have continued to recruit new user projects and collaborators to explore further biomedical areas for applying AI. A number of these projects are built around the communications network facilities we have assembled, bringing together medical and computer science collaborators from remote institutions and making their research programs available to still other remote users. At the same time we have encouraged older mature projects to build their own computing environments thereby freeing up SUMEX resources for newer projects. Nine projects now operate on their own facilities, including three that have become BRTP resources in their own right. Nine projects in the community have completed their research goals and their staffs have moved on to new areas.

- At the end of the current reporting period, we are actively planning the move of the SUMEX and Medical Computer Science offices into newly constructed Stanford Medical School office space, funded by the university. This space, in the Stanford Medical Center complex, provides us with almost twice the area we previously occupied and it is laid out so as to promote better interactions between out groups and among our students and research staff. The move will take place in June.

- SUMEX user projects have made good progress in developing and disseminating effective consultative computer programs for biomedical research. These performance programs provide expertise in analytical biochemical analyses and syntheses, clinical diagnosis and decision-making, molecular biology, and various kinds of cognitive and affective psychological modeling. We have worked hard to meet their needs and are grateful for their expressed appreciation (see Section IV).

III.A.3.2. Resource Equipment Details

The SUMEX-AIM core facility, started in March 1974, was built around a Digital Equipment Corporation (DEC) KI-10 computer and the TENEX operating system which was extended locally to support a dual processor configuration. Because of the operational load on the KI-10's, in the late 1970's, we had added a small DEC 2020 system (see Figure 2) to support more dedicated testing of systems like ONCOCIN and Caduceus and for community demos. This facility provided a superb base for the AI mission of SUMEX-AIM through 1981, when, using DARPA funding, we added a VAX 11/780 system running the UNIX operating system (see Figure 3). By 1982, the KI-10's were becoming difficult to maintain, both in terms of hardware and software, and so were upgraded to a DEC-supported 2060 with about twice the capacity (see Figure 1). The interactive computing environment of this facility, with its AI program development tools and its network and interpersonal communication media, was unsurpassed in other machine environments. Biomedical scientists found SUMEX easy to use in exploring applications of developing artificial intelligence programs for their own work and in stimulating more effective scientific exchanges with colleagues across the country. Coupled through wide-reaching network facilities, these tools also give us access to a large computer science research community, including active artificial intelligence and system development research groups.

The Heterogeneous Computing Environment

In the late 1970's and early 1980's, computer system research on early microprocessors and compact minicomputers suggested that large mainframe computers would not be essential or even the dominant source of computing power for AI research and AI program dissemination. Thus, we began to implement a strategy for computing resources marked by the integration of heterogeneous systems -- mainframes, Lisp workstations, and service systems (e.g., for file storage and printing) all linked together by local area networks. Since the purchase of the first Xerox InterLisp Dolphin workstations in the summer of 1981, many more workstation products have come on the market with significant improvements both in performance and lower cost. Thus, over the years, we have configured the optimal resource computing environment around shared central machines coupled through a high-performance network to growing clusters of personal workstations.

The concept of the individual workstation, especially with the high-bandwidth graphics interface, proved ideal. Both program development tools and facilities for expert system user interactions were substantially improved over what is possible with a central time-shared system. The main shortcomings of early workstation systems were their limited processing speed and high cost. But in the few years since our first experimental systems, processing power has increased by a factor of 10 (e.g., in Texas Instruments Explorers, Symbolics 3600's, and SUN workstations) and cost has decreased by a factor of 3-4 (e.g., in Xerox 1186's). As a concrete overall system example, SUMEX was among the first sites to receive the new Xerox 1186 (DayBreak) last fall. Each one costs less (70% less) than a Dolphin purchased in 1981. Each is roughly 4 times as fast as a Dolphin, has a larger display (20% more pixels), larger disk (38% more space), more memory (245% more), and is small enough and cool enough to fit in a private office or student carrel, eliminating the need for an umbilical cord to a machine-room or floor area therein.

Lisp Workstations

Work in the SUMEX-AIM community and the KSL draws heavily on both of the major dialects of Lisp, Interlisp and the derivatives of MIT's MacLisp. Thus, our workstation purchases have included machines for both environments. We have added new

workstations paced carefully with the developments of higher performing, more compact, and lower cost systems (see Figure 5). By early 1985, we had acquired with NIH funds, DARPA funds, and industrial gifts, the following workstations: 5 Xerox 1100's (Dolphins), 21 Xerox 1108's (Dandelions), 3 Xerox 1109's (DandeTigers), 1 Xerox 1132 (Dorado), 1 Symbolics LM-2, 4 Symbolics 3600's, and 2 Symbolics 3670's.

In late 1985, after long evaluation and vendor negotiations, we acquired a large number of additional workstations including 20 Xerox 1186's (DayBreaks), 20 Texas Instrument Explorers, 1 Symbolics 3600, 1 Symbolics 3640, and 2 Symbolics 3645's. These were purchased mostly with DARPA funding and vendor gifts, and also include two additional drives for the Xerox file server donated last year special software environments for the workstations. These workstations have been broadly integrated into our faculty, staff, and student offices and into public work areas. They are used to support all of our research projects and the gift machines also support various courses offered at Stanford in AI.

We continue to evaluate Lisp workstations as the technology is changing rapidly. Systems based on the SUN Microsystems workstation and the IBM PC-RT have benchmark data rivaling the performance of other specially microcoded Lisp machines (e.g., those from Xerox, Symbolics, and TI), but the software environments are not nearly so extensively developed yet.

Local Area Network Server Hardware

Since the late 1970's, we have been developing a local, high-speed Ethernet environment to provide a flexible basis for planned facility developments and the interconnection of a heterogeneous hardware environment. Our development of Ethernet facilities has been guided by the goals of providing the most effective range of services for SUMEX community needs while remaining compatible with and able to contribute to and draw upon network developments by other groups. We now support primarily 10 Mbit/sec Ethernets (see Figure 6) running numerous protocols and extended geographically throughout the SUMEX-AIM and related Stanford research groups. This network is the "glue" that holds the rest of the computing environment together and consists of numerous servers such as gateways and servers for terminal access, file storage and retrieval, and laser printing.

Hardware for Gateways and TIP's

As we evolved a more complex network topology and decided to compartmentalize the overall Stanford internet to avoid electrical interactions during development and to facilitate different administrative conventions for the use of the various networks, we developed gateways to couple subnetworks together using Motorola MC-68000 systems.

We also developed a MC-68000 terminal interface processor (TIP) to provide terminal access to network hosts and facilities. It is basically a machine that has a number of terminal lines and a network interface and software to manage the establishment of connections for each line and the flow of characters between the terminal and host. It can handle up to 32 lines. Both of these systems are now widely used throughout the Stanford network.

File Server Hardware

Because our Lisp workstations have only limited local file space, the development of effective shared file servers is essential to our resource operation. We had previously implemented two file servers, based on DEC VAX 11/750 machines purchased through a special price arrangement with DARPA (see Figure 4). In the initial file server

configurations, we also bought Fujitsu Eagle 450 MByte disks and controllers (one each from Systems Industries and Emulex) with one 800/1600 BPI tape unit for long term archives, and one 300 Mbyte removable pack drive for cyclic backups. Because of problems with the SI controller, we replaced it this past year with an Emulex so that the two systems are now identical. We also purchased another 450 MByte disk drive to provide needed capacity expansion. Finally, we are in the process of purchasing a SUN 3 file server with two 450 MByte disks using DARPA funds in order to increase file server performance and capacity.

Other Network Hardware

Over this past year, Stanford University has undertaken a major effort to install cabling throughout campus as part of a Stanford University Network (SUNet) development and installation of a new telephone system. These installations have helped improve the connectivity and performance of our network, including redundant links between the new SUMEX and MCS space in the Medical School Office Building and the SUMEX machine room and other parts of campus.
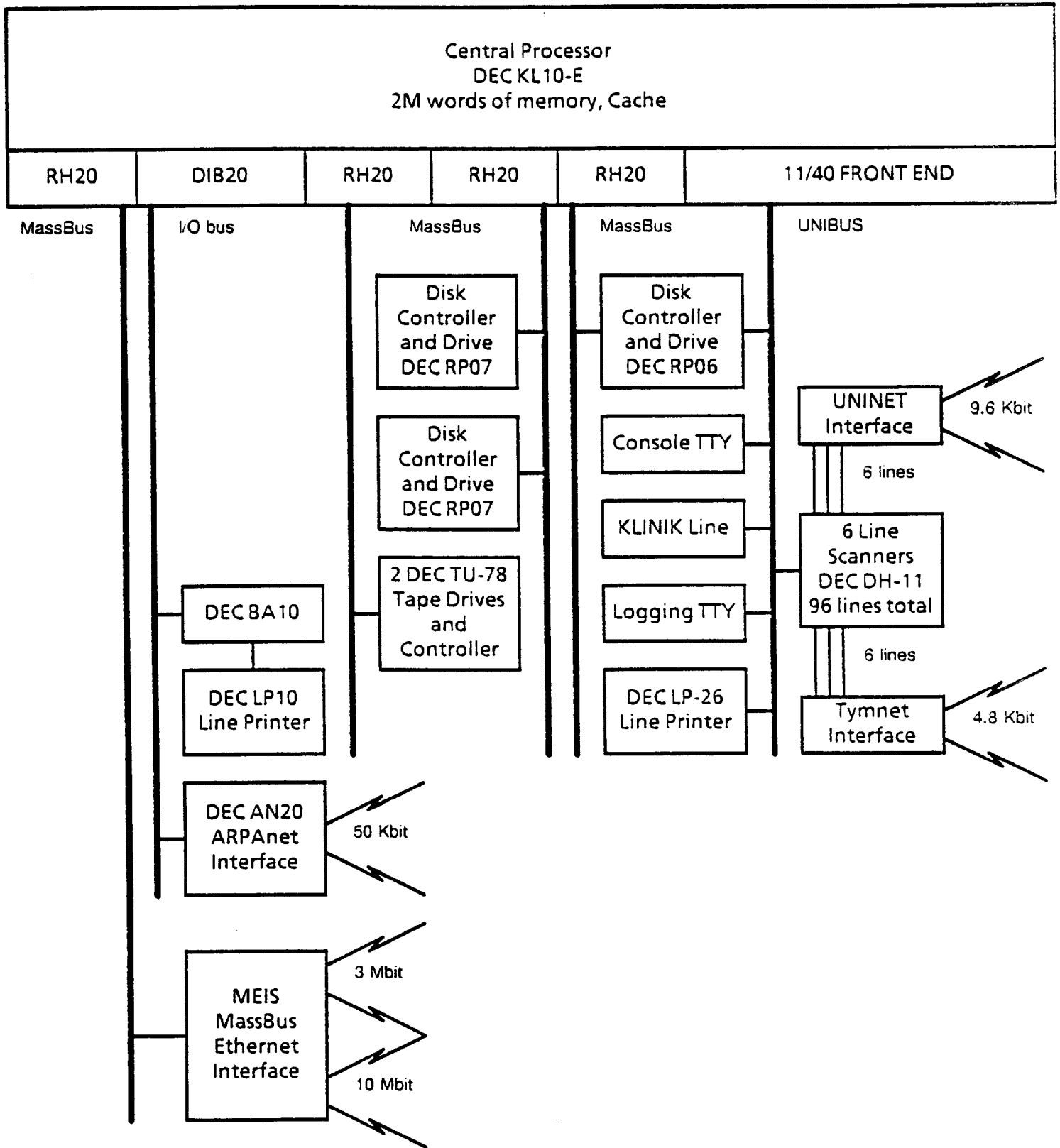
Figure 1: SUMEX-AIM DEC 2060 Configuration

E. H. Shortliffe
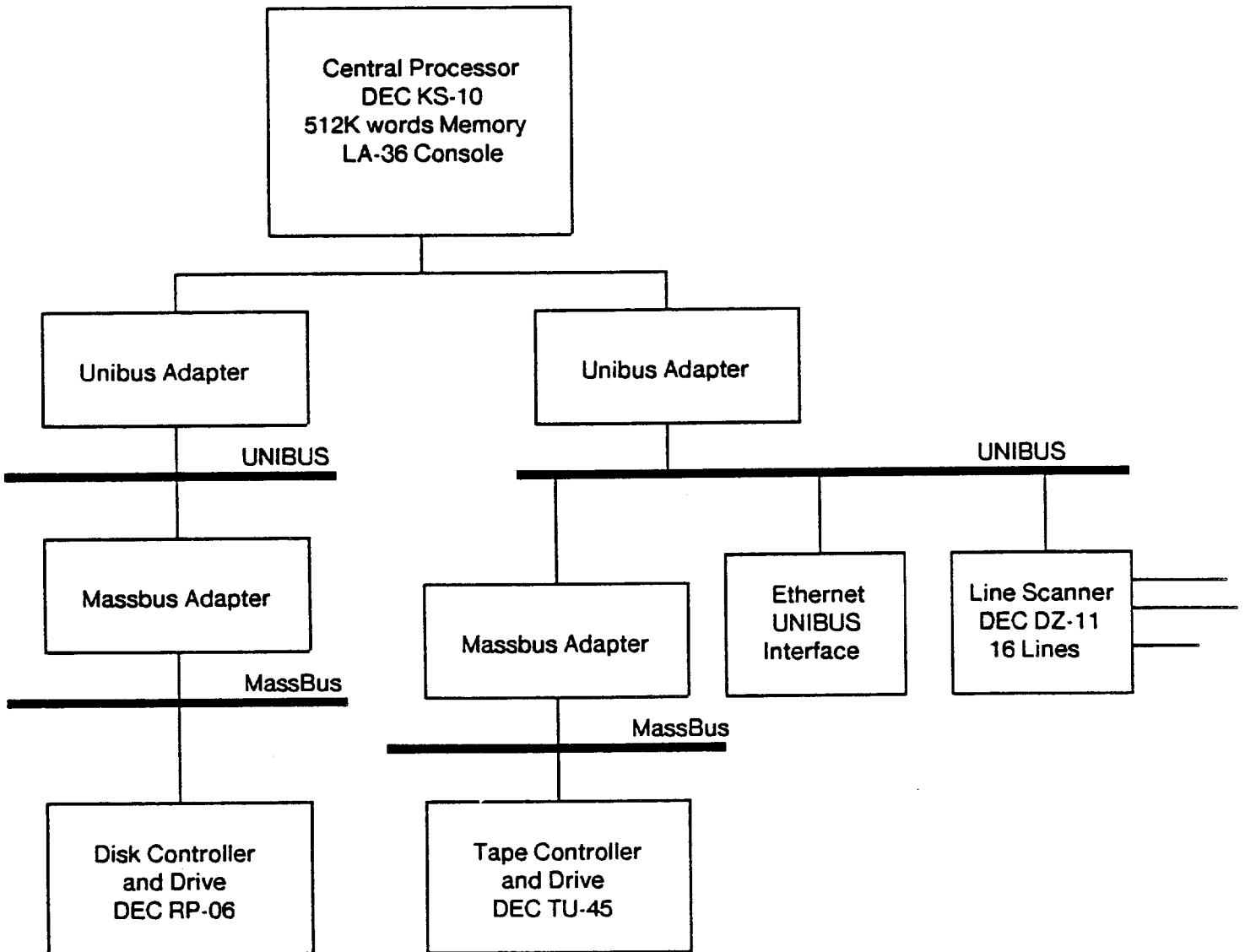
Figure 2: SUMEX-AIM DEC 2020 Configuration
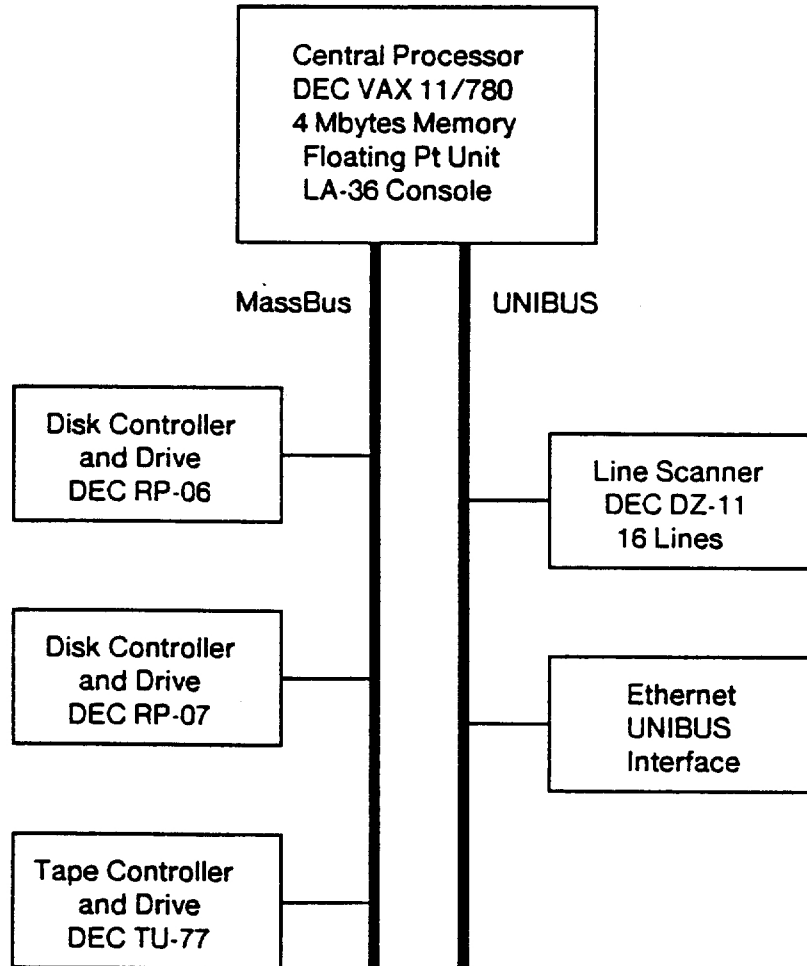
Figure 3: SUMEX-AIM Shared DEC VAX 11/780 Configuration
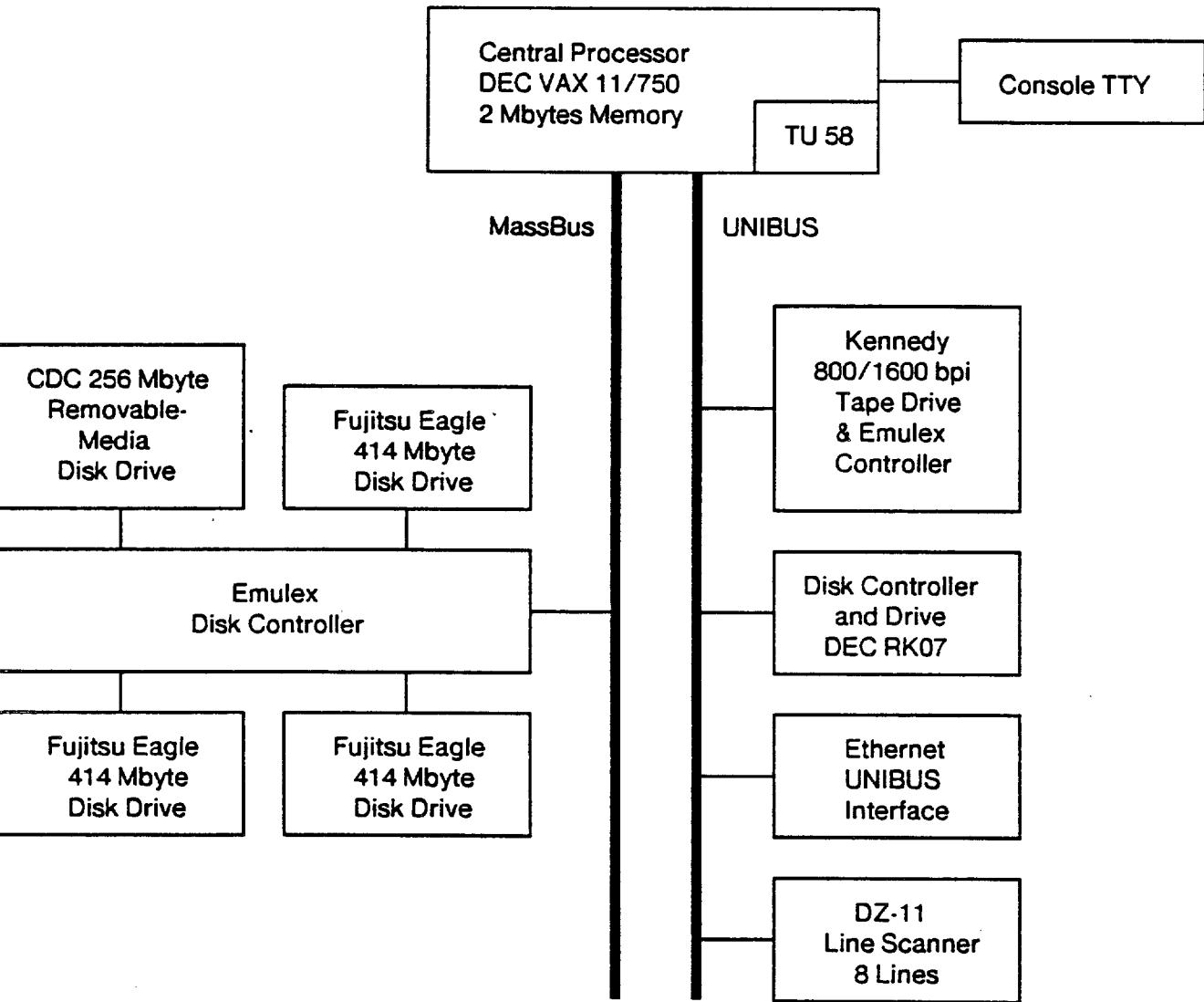
Figure 4: SUMEX-AIM File Server Configuration

Price

200K

100K

X
2060 PSL

X
11/780 PSL

Dorado

Explorer

3600

1100

1109

Lambda

1108

HP PSL

SUN (Lucid)

X 1186

X 1185

1.0

Relative Performance
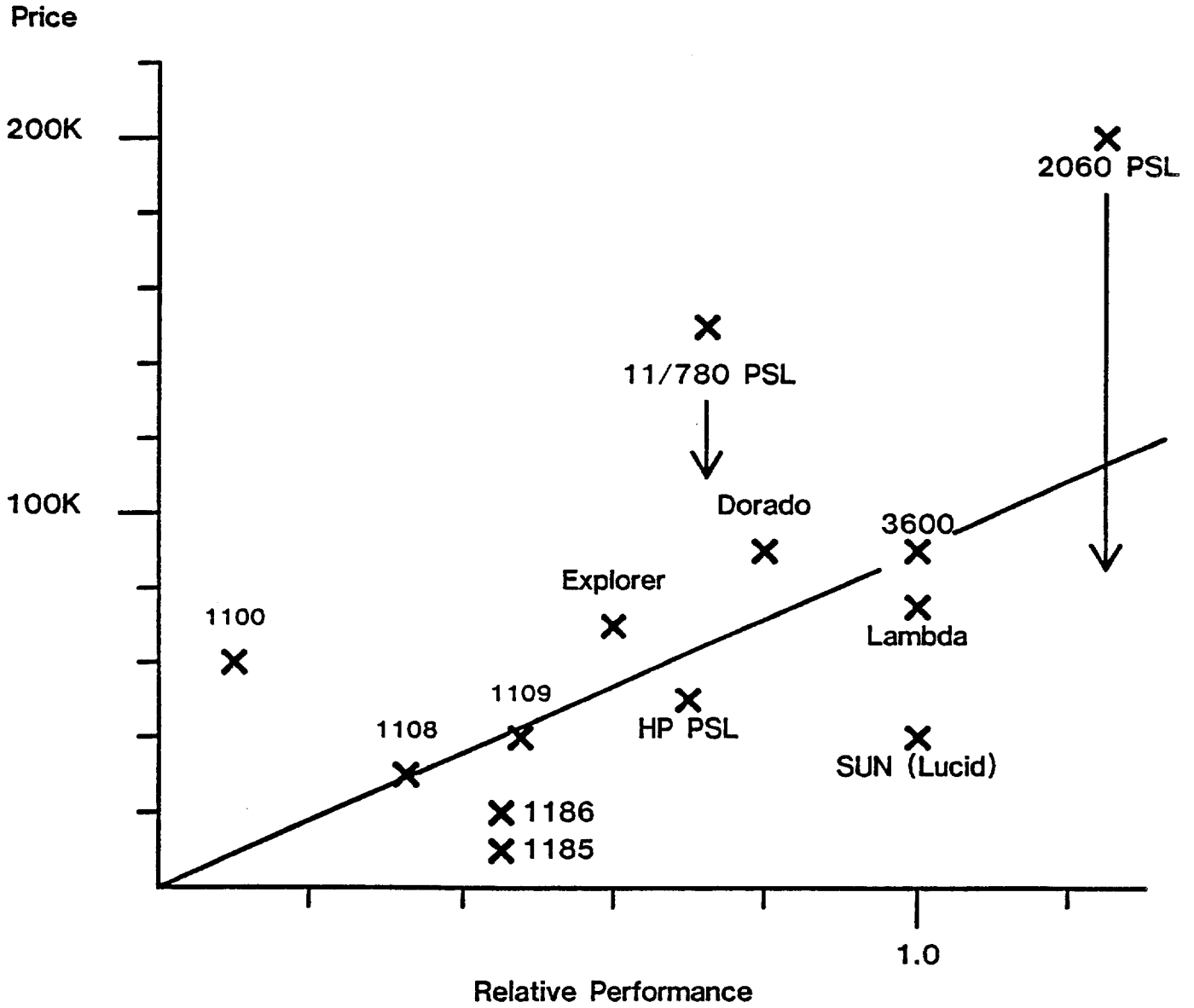
Figure 5:   Price/Performance Comparison of Lisp Workstations

E. H. Shortliffe

Figure 6: SUMEX-AIM EtherNet Configuration

III.A.3.3. Core System Development

Operating System Software

The various parts of the SUMEX-AIM computing environment require development and support of the operating systems that provide the interface between user software and the raw computing capacity. In addition to performance and relevance to AI research, much of our strategy for hardware selection has been based on being able to share development of the operating systems among a large computer science community. This includes the mainframe systems (TOPS-20 and UNIX) and the workstation systems. Following are some highlights of recent system software developments.

TOPS-20 Development and Support

With our long term plan to phase out the 2060 mainframe system, our development efforts in that area are beginning to wind down. Nevertheless, over the past year, considerable work was required to keep the TOPS-20 systems running effectively for the community. This has included the periodic updating, checkout, and installation of new versions of system software. An important upgrade involved moving from the previous 5.3 TOPS-20 monitor release to 6.1. This required a large effort to incorporate all the local SUMEX changes, not only in the monitor itself but in other service software like the EXEC, Galaxy and its spoolers, CHECKD, and ACJ.

Other activities included installation and checkout of the new MCA25 Cache which we purchased the previous grant year to enhance performance by doubling the size of cache memory, the addition of a two-way associative page table, and the addition of a "keep" bit to ret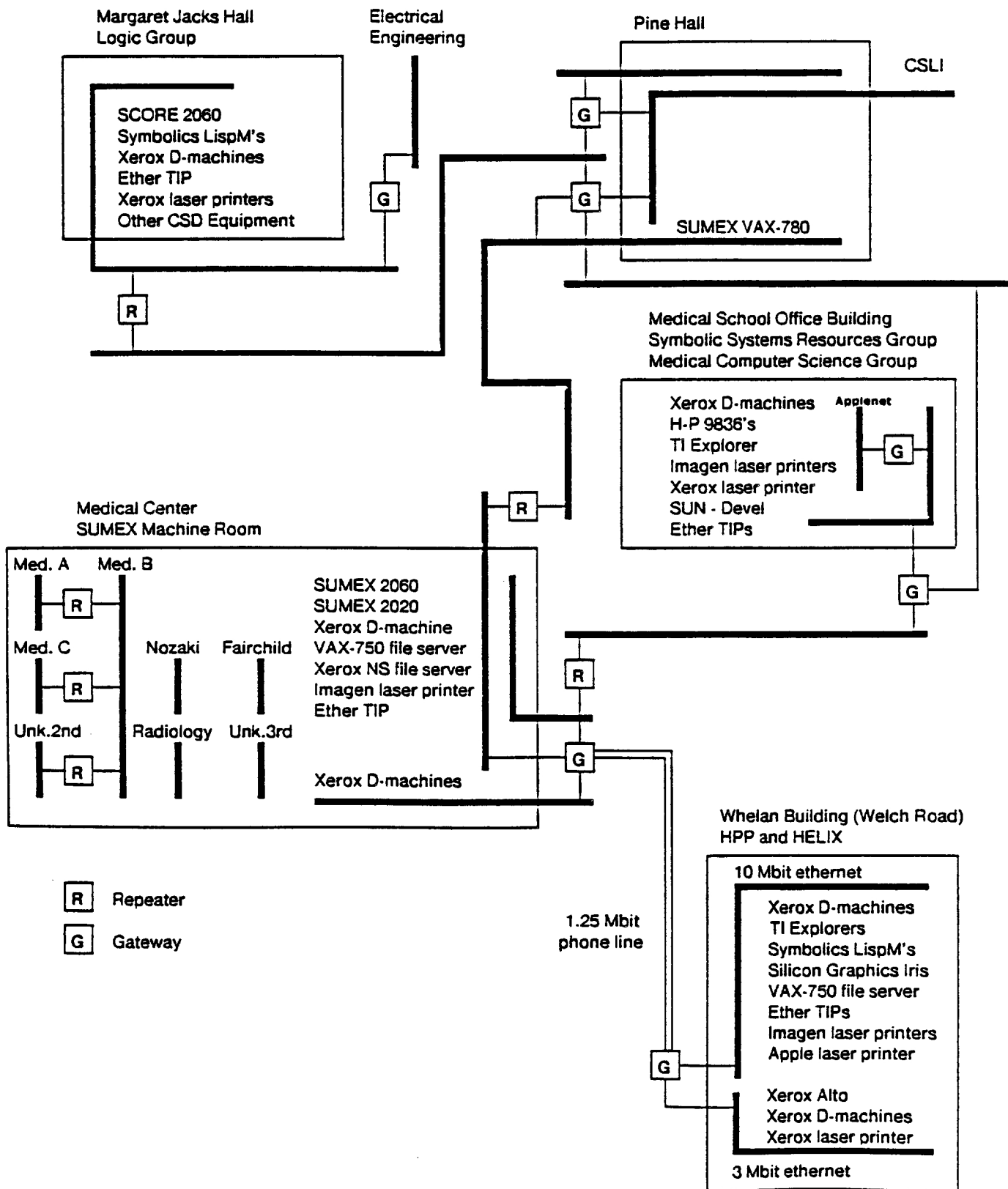ain frequently accessed executive page table entries. Several problems occurred in the installation requiring detailed analysis. SUMEX staff pointed out a bug in DEC's MCA25 diagnostic resulting in a detailed reexamination of the installation which uncovered the omission of one of the backplane wire additions. After correcting the installation, the MCA25 cache worked well as verified by timing tests.

We continued to develop and improve our laser printer spooler, IMPSPL. This included adding many new features such as variable typesizes, page reversing, and manual paper feed for printing on high quality paper. This also required changes to the EXEC and Galaxy to support the new features. IMPSPL appears to be quite solid now, having run for almost a year without any problems.

A significant effort was required "tending the system". In the past year, we analyzed approximately 75 memory dumps of the TOPS-20 operating system in order to pinpoint hardware and software problems and bugs. To facilitate this, we continued the development of QANAL, a Quick automatic dump ANALyze program. Briefly, this program reads the dump and a copy of the monitor file that was running when the failure occurred, and produces a report detailing the state of the system at the time of the failure. Included among these reports is a SYSTAT type output, giving detail to the individual fork, or process, level of all jobs logged in. During the past year, we upgraded QANAL to analyze TOPS-20 release 6.1 data structures, as well as adding several other refinements.

We have continued to track network protocol and service (e.g., file transfer and electronic mail) developments. Many monitor bugs related to TOPS-20 implementation problems relating to the DOD IP/TCP protocols. This complex software required significant effort on our part because SUMEX-AIM has become a major communications crossroads and so exercises the network code very heavily. This has raised many bugs and performance problems that we have worked to improve. We have played an active role in network discussion groups related to areas such as electronic

mail, network designs, and protocols and had kept system tables for network host names and addresses, both local and over the ARPANET, up-to-date.

Because of the heavy loads we have been experiencing on the 2060, we made a detailed performance study of the system dynamics and usage patterns. Among the observations of this study was that the usage of the 2060 was fairly evenly spread between research work (Lisp program development) and text-processing, communications (electronic mail, bulletin boards, etc.), Executive utilities, and system servers (printing, networks, etc.). The research usage has been much higher in the past but has already migrated significantly to personal workstations, while these other applications have not because comparable tools do not exist yet. In this study, we did not find any significant areas of inefficiency in the system -- simply that our user load is very high for the machine resources available.

UNIX Development and Support

We run UNIX on our shared VAX 11/780 and on our 11/750 file servers. This system has been used pretty much as distributed by the University of California at Berkeley, except for local network support modifications, such as for ChaosNet protocols. The local VAX user community is small, so we have not expended much system effort beyond staying current with operating system releases and with useful UNIX community developments.

Workstation System Development and Support

Lisp workstations represent the major new direction for system development at SUMEX-AIM because these machines offer high performance Lisp engines, large address spaces required for sophisticated AI systems, flexible graphics interfaces for users, state-of-the-art program development and debugging tools, and a modularity that promises to be the vehicle for disseminating AI systems into user environments. Accordingly, we have invested a large part of our system effort in developing selected workstations and the related networking environments for effective use in the SUMEX-AIM community. In the transition to workstations as *computing* environments suitable for AI applications work, not just as programming environments, much system development remains to be done, as illustrated below.

Filing

In general, each vendor has addressed the file storage needs of their particular workstation in a way that is incompatible with most other workstations, making support difficult in a highly heterogeneous environment such as the SUMEX-AIM community. The resources necessary to maintain many distinct families of filing conventions and protocols on specialized hardware, all meeting the performance needs of a demanding research community, is prohibitive. Thus we have decided to attempt a compromise.

There is active systems research on distributed file service issues and the results are not clear enough yet to guide long range design decisions. So, we have tentatively decided to adopt a variant of the NFILE file access protocol developed by Symbolics, Inc. A *file access protocol* is intermediate between a *remote file system* and a *file transfer protocol*. A remote file system imposes many constraints upon a potential server machine by specifying features of the file system such as pathname syntax, data block size, character set, protection mechanisms, etc. This makes such protocols very difficult to implement on arbitrary machines as many of these attributes are integrated into operating systems at quite a low level. Conversely, file transfer protocols are specified to allow *copying* an entire file from one machine to another -- a very primitive form of access to the files. A file access protocol can be designed to exist with many

different operating systems, each with its own idiosyncrasies in its file system, but still allowing remote users flexible access to the data stored in the file systems by providing features such as random access, well defined file directory listing, file property manipulation facilities, and asynchronous error recovery.

We decided on NFILE for the following reasons:

- Most significantly, NFILE is built upon abstract interfaces to network streams and host operating systems. It can be easily built upon any reliable byte stream protocol, allowing us to use the results of on-going network development without reimplementing filing protocols. It also is careful not to specify host-dependent information such as pathname syntax or storage format, while providing mechanisms for manipulating file system entities such as directories, files, links, and file attributes. Many file attributes, such as BLOCK-SIZE and CREATION-DATE are included in the protocol, but any can be added as needed, and none are required by the protocol itself, lending flexibility that should make it easy to implement on a variety of operating systems.

- NFILE is a public domain protocol. No licensing is needed to implement or run it.

- An implementation already exists on the Symbolics machine, which can be used for testing and debugging.

- We can implement it to run efficiently in the UNIX kernel, providing the performance the research community needs with inexpensive equipment.

- NFILE can be implemented fairly easily on all of the systems in use by the SUMEX-AIM community since it need not draw on internal operating system features, difficult modifications to existing software should not be needed. Then, as alternative, potentially better techniques become known, NFILE can be abandoned and replaced without having consumed significant resources.

- Many of the options specified by NFILE are derived from the CommonLisp specification and so provide for a significant part of our needs without extension.

Electronic Mail

Electronic mail has become a primary means of communication for the widely spread SUMEX-AIM community. The advent of workstations is forcing a significant rethinking of the mechanisms employed to manage such mail. With mainframes, each user tends to receive and processes mail at the computer he uses most of the time, his *primary host*. The first inclination of many users when an independent workstation is placed in front of them is to begin receiving mail at the workstation, and, in fact, many vendors have implemented facilities to do this. However, this approach has several disadvantages:

- Workstations (especially Lisp workstations) have a software design that gives full control of all aspects of the system to the user at the console. As a result, background tasks, like receiving mail, could well be kept from running for long periods of time either because the user is asking to all of the machine's resources, or because, in the course of working, the user has (perhaps accidentally) manipulated the environment in such a way as to prevent mail reception. This could lead to repeated failed delivery attempts by outside agents.

E. H. Shortliffe

- The hardware failure of a single workstation could keep its user "off the air" for a considerable time since repair of individual workstation units might delayed. Given the growing number of workstations spread throughout office environments, quick repair would not be assured, whereas a centralized mainframe is generally repaired very soon after failure.

- It is more difficult to keep track of mailing addresses when each person is associated with a distinct machine. Consider the difficulty in keeping track of postal addresses or phone numbers if each person you knew lived in a different city. On the other hand, remembering a name and one of several "hosts" is fairly simple, though not perfect.

- It is very difficult to keep a multitude of heterogeneous workstations working properly with complex mailing protocols, making it difficult to move forward as progress is made in electronic communication and as new standards emerge. Each system has to worry about receiving incoming mail, routing and delivering outgoing mail, formatting, storing, and providing for the stability of mailboxes over a variety of possible filing and mailing protocols.

Thus, we are investigating the alternative strategy of having a *mail server* machine which handles *mail transactions*. Because this machine would be isolated from direct user manipulation, it could achieve high software reliability easily, and, as a shared resource, it could achieve high hardware reliability, perhaps through redundancy. The mail server could be used from arbitrary locations, allowing users to be freed from their console to read mail across campus, town, or country without need of expensive machinery.

The mail server acts as an interface among *users, data storage,* and *other mailers.* Users employ a *mail access protocol* to retrieve messages, access and change properties of messages, manage mailboxes, and send mail. This protocol should be simple enough to implement on relatively simple, inexpensive machines so that mail can be read remotely easily. This is somewhat distinct from some previous approaches since the mail access protocol is used for all message manipulations, isolating the user from all knowledge of how the data storage is used. This means the the mail server can utilize the data storage in whatever way is most efficient to organize the mail. The data storage could be anything from conventional magnetic disk file system to a highly specialized mail filing system built on optical disks, since it is abstracted from other elements in the mail system. The other mailers constitute the mail server's (and thus the users') link to the outside world. The mail server would use various *mail transport protocols* (e.g., SMTP) to exchange mail with other mail hosts.

We have been investigating user mail interface issues for workstations, as well as issues for the mail access protocol itself. We are examining several related projects, including MIT's PCMAIL, the public parts of Xerox's Grapevine and NSMail, and work on Stanford's V system. We have implemented an interim mail access protocol and have begun implementing user interfaces that make use of it on Xerox D-machines and Texas Instruments Explorers.

Xerox D-Machines

Much of the SUMEX-AIM community uses InterLisp and has moved naturally to the Xerox D-machines -- initially the Dolphin (1100), then the Dandelion (1108), Dandetiger (1109), and Dorado (1132), and now the DayBreak (1186). Much work has gone into hardware installation and networking support but we have also developed numerous software packages to help make the machines more effective for users and to ease our own problems in managing the distributed workstation environment.

The number and utility of "lispusers packages" has again increased significantly over the past year. Although too numerous to detail (there are approximately 550 packages currently, up from about 240 last year), packages receiving heavy use for the first time in the past year were Sketch, FileBrowser, TEdit, Manager, Impress, Hash, Helpsys, and Spy. Many of these packages were beta-tested at SUMEX and/or patches and updates were distributed via the Info-1100 and Bug-1100 discussion lists we maintain. Other AIM sites and research groups around the world were able to share in our progress and benefit from our experience by participating in the discussion lists. We, in turn, as subscribers to the same discussion lists were able to benefit from the experiences and expertise of others.

The past year saw Interlisp's device-independent graphics mechanism ("image streams") mature and grow a good deal. We participated in the design changes, ensuring that the specification was sufficiently device-independent that our Impress laser printer protocol package would be integrated into the system as easily and well as the Xerox-authored Press protocol and Interpress protocol packages. An ImageStream driver was developed to support a Hewlett-Packard color plotter. The development of the driver helped to explore the issues of color in the ImageStream specification as well as test how it applied to analog devices. For the first time we were able to generate color hardcopy output from the both the black & white Interlisp workstations and, with some additional conversion software, from the color Iris workstation.

The Impress package was extended to include almost all of the operators in the new specification. The Impress package is sufficiently complete that HARDCOPYW, DISPLAYGRAPH, TEdit, and Sketch all produce output of quality comparable to that of the more-expensive Xerox laser printers. In most cases the output is generated faster than for those printers and more compactly. We are presently participating in another round of improvements to the specification.

In support of the expanding number of ImageStream drivers, a self-scaling graphics command set was implemented on top of the standard graphics command set that allowed software to be written without regard to the scaling requirements of a particular output device. This allowed graphics output to be easily redirected to varying printing devices without modification of the source program and/or without adding additional scaling routines to every program.

Implementation of an Interlisp-based Ethernet boot file server for the Xerox workstations was completed this past year. This server made it possible to obtain workstation installation and diagnostic utilities via the network as an alternative to floppy disks. Much later, the Interlisp-based boot file server was replaced with a Xerox product Ethernet boot server which extended our network installation and diagnostic capabilities. The addition of network boot file service has led to improvements in software installation procedures by allowing us to move away from our previous dependence on floppy disks. This has become increasingly important as the newer Xerox 1186 hardware supports a smaller capacity floppy disk drive and would require use of over a dozen floppy disks if Ethernet installation were not available.

Initial exploration into distributed systems and remote workstation access was started. An experimental XNS-based TELNET server was built to allow access into a workstation remotely via the Ethernet. This experimental server uncovered numerous problems with the workstation software and initiated discussions that led to a complete TELNET/GAP (Xerox XNS Gateway Access Protocol) server for the workstations (pending the next Xerox software release). The workstation "executive" part of the experimental TELNET server was extracted and generalized and put to use in the TCP-based Ethernet virtual graphics work.

We developed a system called IMEDIT. This program allows users to break apart Impress files and also to merge in other Impress files. Merging Impress files is an

important feature since SCRIBE cannot do .this. SCRIBE can merge in a picture but any text in the picture will be printed in the wrong font. IMEDIT gets around this problem by manipulating the fonts so that fonts in a merged file don't conflict with fonts in the base file. IMEDIT can also generate an ASCII file showing the Impress commands and their arguments in an Impress file. This feature is invaluable for those who need to understand the Impress language.

Currently we are working on the TEdit text editor, initially to facilitate simple document types like memos. We have implemented an ImageObject that allows users to select the logo they prefer and are working on others for document features like the return address. Eventually users will be able to interactively choose what they want from standard menus. Such systems are essential to allow users to move work from the 2060 to workstations.

We have worked closely with many other sites, including the Center for Study of Language and Information at Stanford, the Stanford Campus Networking group, Rutgers University, Ohio State University, the University of Pittsburgh, Cornell, Maryland, and industrial research groups such as Xerox Palo Alto Research Center, SRI, Teknowledge, IntelliCorp, and Schlumberger-Doll Research. We have been the maintainers for the international electronic mail network of users for research D-machines, which have upwards of 300 readers, and the interchange of ideas and problems among this group has been of great service to all users.

ZetaLisp Workstations

The complement of ZetaLisp-based workstations has grown to include twenty Texas Instruments Explorers and ten Symbolics 3600-class machines. The acquisition of these machines was driven by three primary factors. First, many of the research projects are attempting to become independent of any particular machine, and so are moving development to the CommonLisp standard language. These machines are among the first to offer production quality support of CommonLisp. Second, some application systems require substantial performance in terms of processing speed and address space in order to complete in a reasonable amount of time. These machines were among the highest performing Lisp machines available at the time. Finally, there are researchers who prefer the MacLisp/Emacs derived programming environment.

The Explorer and the 3600 are both built on MIT's ZetaLisp software, and so continue to share much functionality. Therefore, many projects have been undertaken simultaneously on both machines. In order to facilitate this interoperability, two compatibility packages have been built, one for each type of machine. The packages contain code to add functionality to each machine to bring it closer to the specification of the other machine where possible, without blocking the native functionality of the system it is running on. There are also lists of features which do not exist and could not be easily duplicated, as well as suggested workarounds where appropriate. These compatibility packages have been made available to the ARPANET community.

We found that users of these machines spent a considerable amount of time redoing work that had already been done since there was no adequate library of user-written tools to draw from. Thus we have undertaken to provide such a facility and to gather as many tools as possible. The *TOOLS* system allows a user to select those tools that he wishes to load either by giving a list of their names (in an initialization file, for instance), or by selecting them from a menu. The menu can also be used to obtain on-line documentation about each tool, and so provides a convenient way to browse the tools. The following is a list of the tools that have so far been implemented or collected:

FS-TO-FS-BACKUP -- Functions that can copy unbacked-up files from the Lisp

machine's file system to another file system which is then backed up to tape. This obviates the need to do backups to expensive and slow cartridge tapes.

SYSTEM-MANAGER -- Provides for shared access to hierarchically structured sets of files, allowing multiple people to work on development of a single system or subsystem simultaneously.

DYNAMIC-SYSTEM-MENU -- Attempts to facilitate managing the screen so that the size and position of windows can be easily tailored to the task at hand.

NET-IMAGEN -- Allows printing on network based Imagen printers using the Impress document formatting language. (Symbolics implementation from MIT)

TCP-FINGER -- (Needed only on Explorer) Implements the popular FINGER person lookup protocol for TCP/IP.

WHO-LINE -- (Explorer only) Shows percentage-wise progress through editor buffers during lengthy operations such as compilation.

VERTICALLY-ORDERED-MENU-ITEMS -- Allows multi-column menus to be displayed with the items split into columns first rather than rows first.

SMALL-FONTS -- Changes all standard windows to use a smaller font, allowing more data to be displayed at a time.

SCREEN-EDIT-MIXIN -- Allows selected windows to be moved or reshaped by clicking on small boxes in the margins of the windows.

MOUSE-SELECTABLE-PANE-MIXIN -- Allows constraint frame panes to be mouse selectable.

MAKE-INTO-SCRIBE-FILE -- Converts a file with Lisp machine special characters into Scribe format so that the special characters will be correctly printed on Imagen printers.

INSPECT-HASH-TABLES -- (Needed on Explorer only) Causes the inspector to display hashing data structures in a more readable Key/Value format.

GENERAL-NAMED-STRUCTURE-MESSAGE-HANDLER -- Causes selected structures to error instead of returning NIL when they receive messages they do not handle, facilitating debugging.

FILTER-WINDOW-DEBUGGER -- Allows specifying functions that will not be displayed in the window debugger, eliminating the clutter of system functions so that user is only presented with "interesting" stack frames.

DEFSTRUCT-TYPE-CHECKING -- An addition to DEFSTRUCT that causes the access functions of selected structures to check their arguments, facilitating debugging.

DEBUG-STACK-GROUP -- (Explorer only) Allows users entering the debugger to examine a particular stack group from the window debugger.

BATCH-PROCESSOR -- Facility for "running" command files overnight.

CHOOSE-VARIABLE-VALUES-MACROS -- Alternate interface to the CHOOSE-

> VARIABLE-VALUES facility which does not require the user to specify and manipulate specials.

All of these tools except for parts of NET-IMAGEN were developed at Stanford. We are currently working with the Lisp machine vendors on licensing that will make it possible to distribute these tools via the ARPANET. (Both Symbolics and TI are much more restrictive in their software-sharing policies than is Xerox).

A great deal of work has been done in installing the Explorers in the SUMEX-AIM research environment. This is one of the first installations of a large number of Explorers, and we have participated actively in the "shaking down" of the Explorer, by being a beta test site for release 2 of the Explorer system software, and release 1 of the Explorer TCP/IP software. In the course of the testing, we submitted 56 written software problem reports, and over 25 verbal reports, many of which included solutions. As a result, the Explorer is now a well-integrated part of the research environment, allowing many researchers to actively pursue their work by putting powerful development tools on their desks.

Virtual Workstation Graphics

We have done a number of experiments with the remote connection of bitmapped displays to hosts and workstations. Generally, the displays on Lisp machines are tethered through a high bandwidth cable to their processors. This limits the flexibility with which users can move from one Lisp machine to another (one must move physically to another machine) and loses the ability of researchers to work from home over telephone lines. A way of providing a more flexible display to processor connection is to use a virtual graphics protocol, such as the V Kernel system developed by Lantz [4]. This allows efficient communication of the contents of a display window to be compactly represented, transmitted over a communication network, and reconstructed on a remote bitmapped screen.

In order to more fully understand the integration of remote virtual graphics access to workstation, a nearly complete implementation of a client interface of the Lantz Virtual Graphics Protocol (VGP) was done within the Xerox Interlisp-D environment. This implementation was done in such a way as to make the fact that the VGP was in the system transparent to the Interlisp programmer. Several key steps were involved:

- Since the access to the workstation was to be done remotely on some kind of network, it was necessary to write an IP/TCP/TELNET server which handled the peculiarities of the TELNET protocol, and provided the usual input and output data streams to a virtual graphics stream executive. This executive then did remote login authentification and called the remote LISP evaluator.

- In parallel to the LISP evaluation, a mechanism was necessary to interface the client VGP into Interlisp-D. This was done using the IMAGEOP objects which are associated with each stream within Interlisp-D. As a consequence all reads and writes on the "standard" input and output streams were defaulted to the VGP input and output streams when the workstation was accessed remotely in this manner.

- As a consequence when one is connected to such a workstation through a VGP server, the graphics engine was driven by standard calls to graphics functions on the Interlisp-D workstation. Thus, the functionality of windows, menus, text within windows, mouse interaction, and a suite of drawing functions were all translated to the VGP and done remotely on the user's workstation.

This feasibility experiment proved that remote access to a LISP workstation using virtual graphics protocols was practical. This paves the way for additional work to allow researchers to take advantage of powerful user/graphics environments on Lisp machines, even if not physically near the machine.

Network Services

A highly important aspect of the SUMEX system is effective communication within our growing distributed computing environment and with remote users. In addition to the economic arguments for terminal access, networking offers other advantages for shared computing. These include improved inter-user communications, more effective software sharing, uniform user access to multiple machines and special purpose resources, convenient file transfers, more effective backup, and co-processing between remote machines. Networks are crucial for maintaining the collaborative scientific and software contacts within the SUMEX-AIM community.

Remote Networks

In addition to continuing our connection to TYMNET, we have implemented an experimental connection to UNINET this past year in an attempt to improve services for remote users. As reported last time, we have had serious difficulties getting needed service from TYMNET for debugging network problems and users away from major cities have problems with echo response times. The opinions of about 15 of our heaviest TYMNET users were sought concerning the performance of TYMNET. Though many were quite pleased, several with experience on a variety of other such networks recommended a change. The TYMNET hardware interface itself has been quite dependable in the past year.

Discussions were held with CompuServe and UNINET concerning alternative service. The UNINET connection was finally installed after a period of considerable review, based on technical evaluations, cost analyses, and the experience of other network customers with similar systems (e.g., BIONET). Requests for office phone locations went out to 70 of our TYMNET users. Responses from 50 of them revealed only one who would be unable to reach a UNINET node with a local phone call. The capability of KERMIT data transfers was reviewed as was the capability for using the text editors, EMACS and TVEDIT. On the whole, UNINET seems much more responsive for current users, although evaluation is still underway. Both TYMNET and UNINET services are purchased jointly with the Rutgers Computers in Biomedicine resource to maximize our volume usage price break.

We also continue our extremely advantageous connection to the Department of Defense's ARPANET, managed by the Defense Communications Agency (DCA). This connection has been possible because of the long-standing basic research effort in AI within the Knowledge Systems Laboratory that is funded by DARPA. ARPANET is the primary link between SUMEX and other machine resources such as Rutgers-AIM and the large AI computer science community supported by DARPA. We are also attempting to establish a link to the DARPA wideband satellite network to facilitate the rapid transfer of large amounts of data such as are involved with projects like our Concurrent Symbolic Computing Architectures project.

Local Area Networks

For many years now, we have been developing our local area networking systems to enhance the facilities available to researchers. Much of this work has centered on the effective integration of distributed computing resources in the form of mainframes, workstations, and servers. Network gateways and terminal interface processors (TIP's)

were developed and extended to link our environment together and are now the standard system used in the campus-wide Stanford University network. We are developing gateways to interface other equipment as needed too. A diagram of our local area network system is shown in Figure 6 and the following summarizes our LAN-related development work.

Ethernet Gateways -- In our heterogeneous network environment, in order to provide workstation access to file servers, mail servers, and other computers within the network, it is necessary to able to route multiple networking protocols through the network gateways. Over the past year, support for both the Xerox NS and Symbolics/Texas-Instrument CHAOSNET protocols were added to the SUMEX gateways. This support not only provides the routers necessary to move such packets within this topology, but also other miscellaneous services such as time, name/address lookup, host statistics, address resolution, and routing table broadcast and query information. As a consequence, the SUMEX gateways now support these protocols as well as the PUP, and IP protocols. These services are unique within the SUMEX-AIM portion of the Stanford University network, and give our researchers a networking environment that is flexible, of high bandwidth, and extremely dependable.

Remote Ethernet -- Some preliminary design was done on a "home Ethernet connection" to facilitate virtual graphics access and other network connections from home workstations. The feasibility of this device was investigated to evaluate its cost versus a similar device manufactured by Bridge Systems at a cost of $5,000.00. It is believed that a less expensive device can be built that will conform to our remote Ethernet needs. Although this device will communicate via modems and hence be much slower than the 10 MBit/sec Ethernet bandwidth, the fact that each remote station will act as an Ethernet host without the RS-232 overhead, should improve file transfer significantly.

Network Bootstrap -- Over the past year, SUMEX has participated in the definition of a remote workstation bootstrap protocol which can flexibly load systems over networks, even through gateway links to remote servers. The details of the protocol are documented in RFC951, put out by the ARPANET development group. Implementation of the BOOTP protocols required developing a new programmable read-only memory (PROM) monitor for our workstations (MC 68000-based) with a more extensive command structure to facilitate specification of remote boot file pathnames. If the user specifies enough information (server address, workstation address, and file name), then the PROM bypasses the BOOTP phase entirely and directly enters the transparent FTP phase. This can be useful for manually booting from arbitrary internet hosts not running BOOTP servers. The PROM code currently contains drivers for the 3COM 3C400 interface (at 4 possible multibus board addresses) and the Interlan NI3210 (also at 4 addresses).

The BOOTP/TFTP bootstrap uses a global structure located at the end of memory during its operation. This structure is left intact after the booted program gets control. In some cases a program (such as an EtherTIP) may want to fetch a configuration file listing its addresses and options before starting up. With the mechanism provided by this structure, that program can call the PROM resident TFTP code to fetch the desired configuration file.

Network gateway modifications necessary to route BOOTP requests have been made and installed in all Stanford gateways.

Laser Printing Services

Since the first Xerox laser printers were developed in the mid-1970's, a number of companies have produced computer-driven systems, such as Imagen and Adobe. These systems have become essential components of the work environment of the SUMEX-AIM community with applications ranging from scientific publications to hardcopy graphics output for ONCOCIN chemotherapy protocol patient charts. We have done much systems work to integrate laser printers into the SUMEX network environment so they would be routinely accessible from hosts and workstations alike.

Over the past year, we purchased 2 new Imagen 12/300's, upgraded an 8/300 to a 12/300, and converted an old Hewlett-Packard 2688A to a 12/300 laser printer for the SUMEX-AIM community. These enhancements were funded by DARPA. The move to 12/300's was motivated primarily by the ruggedness of the Ricoh LP-4120 print engine used in those printers. Whereas the Canon LBP-CX print engine used in the 8/300 has an expected lifetime of 70,000 pages, the Ricoh LP-4120 has an expected lifetime of 700,000 pages. Since the KSL printed roughly 250,000 pages on laser printers last year we decided it was time to move to a sturdier printing workhorse. Other beneficial side-effects of the upgrade were: (1) higher print rate (12 pages-per-minute), (2) bigger paper tray (half a ream), (3) blacker and more solid print, (4) crisper print, and (5) cheaper supplies (half the price per page compared to the 8/300).

We have also acquired an Apple Laser Writer which interprets the PostScript page description language. Within a few months of its introduction, the Apple Laser Writer has become the most common laser printer on campus and around the world. Economies of scale have made it possible for us to acquire this printer for under $4000. SUMEX AppleNet/Ethernet expertise will make it possible for us to attach the Laser Writer to the high-bandwidth campus internet and operate the printer at the high-end of its 8 page-per-minute capacity. (The vast majority of laboratory-owned Laser Writers in the U.S. are driven over a low-bandwidth RS-232 line yielding only 3 pages-per-minute throughput and typically greater latency.) The PostScript page description language is already the standard of choice at university and DARPA sites (judging by traffic on the Laser-Lovers discussion group). It is generally agreed upon in these communities that PostScript is among the easiest-to-generate and most expressive of the page description languages in use today and reconciles these traits much more effectively than other languages do.

Although anyone with $4000 can benefit from the advantages of owning a Laser Writer, SUMEX users at Stanford have access to a Linotronics 300P typesetter owned by the university. This printer interprets PostScript files identical to those which can be printed on a Laser Writer, but renders its output on photographic paper up to 11" x 17" in size at a resolution of 1200 scans-per-inch. (At present, most of our printers image at 300 spi and our finest printer is the aging Xerox Alto-Raven which images at 384 spi.) To exploit the special capabilities of this printer and to take advantage of the economical Apple Laser Writer, we have begun an Interlisp implementation of an "image stream" driver for PostScript. Unilogic has already added Postscript support to Scribe and Adobe has implemented Postscript support for TeX.

General User Software

We have continued to assemble (develop where necessary) and maintain a broad range of user support software. These include such tools as language systems, statistics packages, vendor-supplied programs, text editors, text search programs, file space management programs, graphics support, a batch program execution monitor, text formatting and justification assistance, magnetic tape conversion aids, and user information/help assistance programs.

E. H. Shortliffe

A particularly important area of user software for our community effort is a set of tools for inter-user communications. We have built up a group of programs to facilitate many aspects of communications including interpersonal electronic mail, a "bulletin board" system for various special interest groups to bridge the gap between private mail and formal system documents, and tools for terminal connections and file transfers between SUMEX and various external hosts. Examples of work on these sorts of programs have already been mentioned in earlier sections on operating systems and networking.

At SUMEX-AIM we are committed to importing rather than reinventing software where possible. As noted above, a number of the packages we have brought up are from outside groups. Many avenues exist for sharing between the system staff, various user projects, other facilities, and vendors. The availability of fast and convenient communication facilities coupling communities of computer facilities has made possible effective intergroup cooperation and decentralized maintenance of software packages. The many operating system and system software interest groups (e.g., TOPS-20, UNIX, D-Machines, network protocols, etc.) that have grown up by means of the ARPANET have been a good model for this kind of exchange. The other major advantage is that as a by-product of the constant communication about particular software, personal connections between staff members of the various sites develop. These connections serve to pass general information about software tools and to encourage the exchange of ideas among the sites and even vendors as appropriate to our research mission. We continue to import significant amounts of system software from other ARPANET sites, reciprocating with our own local developments. Interactions have included mutual backup support, experience with various hardware configurations, experience with new types of computers and operating systems, designs for local networks, operating system enhancements, utility or language software, and user project collaborations. We have assisted groups that have interacted with SUMEX user projects get access to software available in our community (for more details, see the section on Dissemination on page 89).

Operations and Support

The diverse computing environment that SUMEX-AIM provides requires a significant effort at operations and support to keep the resource responsive to community project needs. This includes the planning and management of physical facilities such as machine rooms and communications, system operations routine to backup and retrieve user files in a timely manner, and user support for communications, systems, and software advice. Of course, the move of our groups to new space in the Medical School Office Building has required major planning and care to ensure minimum downtime for our computing environment and much systems and electronics work to outfit the new space.

We use students for much of our operations and related systems programming work. We spend significant time on new product review and evaluation such as Lisp workstations, terminals, communications equipment, network equipment, microprocessor systems, mainframe developments, and peripheral equipment. We also pay close attention to available video production and projection equipment, which has proved so useful in our dissemination efforts involving video tapes of our work.

III.A.3.4. Core AI Research

We have maintained a strong core AI research effort in the SUMEX-AIM resource aimed at developing information resources, basic AI research, and tools of general interest to the SUMEX-AIM community. It should be noted that the SUMEX resource grant from NIH supports much of the computing environment for this core AI work[1] but NIH supports only a small part of the manpower and other support for core AI. Substantial additional support for the personnel costs of our core AI research (roughly comparable to the NIH investment in computing resources) comes from DARPA, ONR, NSF, NASA, and several industrial basic research contracts to the Knowledge Systems Laboratory.

The following summary reports progress on the basic or core research activities within the KSL. The development of the ONCOCIN system (under Professor Shortliffe) is an important part of our core research proposal for the renewal period. Progress on that work is reported separately in Section IV.A.3, however, because its efforts have been supported as a collaborative and resource-related research project up until now. Together, this work explores a broad range of basic research ideas in many application settings, all of which contribute in the long term to improved knowledge based systems in biomedicine.

Rationale

Our core AI research work has long been the mainstay on which our extensive list of applications projects are based. Medical information -- both medical data and medical knowledge -- is the key to progress in research and excellence in biomedical science and clinical practice. As the rapid explosion of information continues, clinicians and biomedical scientists must turn to computers for help in managing the information, and in applying it to complex situations.

Artificial Intelligence (AI) methods are particularly appropriate for aiding in the management and application of knowledge because they apply to information represented symbolically, as well as numerically, and to reasoning with judgmental rules as well as logical ones. They have been focused on medical and biological problems for over two decades with considerable success. This is because, of all the computing methods known, AI methods are the only ones that deal explicitly with symbolic information and problem solving and with knowledge that is heuristic (experiential) as well as factual.

Expert systems are one important class of applications of AI to complex problems -- in medicine, science, engineering, and elsewhere. An expert system is one whose performance level rivals that of a human expert because it has extensive domain knowledge (usually derived from a human expert); it can reason about its knowledge to solve difficult problems in the domain; it can explain its line of reasoning much as a human expert can; and it is flexible enough to incorporate new knowledge without reprogramming. Expert Systems draw on the current stock of ideas in AI, for example, about representing and using knowledge. They are adequate for capturing problem-solving expertise for many bounded problem areas. Numerous high-performance, expert systems have resulted from this work in such diverse fields as analytical chemistry, medical diagnosis, cancer chemotherapy management, VLSI design, machine fault diagnosis, and molecular biology. Some of these programs rival human experts in solving problems in particular domains and some are being adapted for commercial use.

---

[1]DARPA funds have also helped substantially in upgrading our mainframe systems and in the purchase of community Lisp workstations

Other core research projects have developed generalized software tools for representing and utilizing knowledge (e.g., EMYCIN, UNITS, AGE, MRS, GLISP) as well as comprehensive publications such as the three-volume *Handbook of Artificial Intelligence* and books summarizing lessons learned in the DENDRAL and MYCIN research projects.

There is considerable power in the current stock of techniques, as exemplified by the rate of transfer of ideas from the research laboratory to commercial practice. But it is also clear that today's technology needs to be augmented to deal with the complexity of medical information processing.

Our core research goals, as outlined in the next section, are to analyze the limitations of current techniques and to investigate the nature of methods for overcoming them. Long-term success of computer-based aids in medicine and biology depend on improving the programming methods available for representing and using domain knowledge. That knowledge is inherently complex: it contains mixtures of symbolic and numeric facts and relations, many of them uncertain; it contains knowledge at different levels of abstraction and in seemingly inconsistent frameworks; and it links examples and exception clauses with rules of thumb as well as with theoretical principles. Current techniques have been successful only insofar as they severely limit this complexity. As the applications become more far-reaching, computer programs will have to deal more effectively with richer expressions and much more voluminous amounts of knowledge.

Highlights of Progress

In the last year, research has progressed on several fundamental issues of AI. As in the past, our research methodology is experimental; we believe it is most fruitful at this stage of AI research to raise questions, examine issues, and test hypotheses in the context of specific problems such as management of patients with Hodgkin's disease. Thus, within the KSL we build systems that implement our ideas for answering (or shedding some light on) fundamental questions; we experiment with those systems to determine the strengths and limits of the ideas; we redesign and test more; we attempt to generalize the ideas from the domain of implementation to other domains; and we publish details of the experiments. Many of these specific problem domains are medical or biological. In this way we believe the KSL has made substantial contributions to core research problems of interest not just to the AIM community but to AI in general.

In addition to the technical reports listed below, the following survey articles were published during this year. These are of central interest to AI researchers and of direct relevance to the mission of the SUMEX-AIM resource.

SURVEY ARTICLES: KSL-85-19, KSL-85-27, KSL-85-28, KSL-85-37, KSL-85-54, KSL-86-17, KSL-86-32

Progress is reported below under each of the major topics of our work. Citations are to KSL technical reports listed in the publications section.

1. *Knowledge representation:* How can the knowledge necessary for complex problem solving be represented for its most effective use in automatic inference processes? Often, the knowledge obtained from experts is heuristic knowledge, gained from many years of experience. How can this knowledge, with its inherent vagueness and uncertainty, be represented and applied?

   Work continues on BB1, with its explicit representation of control knowledge, as reported last year. In addition, part of our research on

NEOMYCIN is focused on using a flexible, rich representation of control knowledge so that we can model problem solving at the strategic level as well as at the tactical level.

[See KSL technical memos KSL-85-16, KSL-85-17, KSL-85-31, KSL-86-11, KSL-86-27.]

2. *Blackboard Architectures and Control:* How can we design flexible control structures for powerful problem solving programs?

We have continued to develop the BB1 blackboard architecture for systems that reason about -- control, explain, and learn about -- their own actions. We have developed domain-independent control knowledge sources for refining abstract control plans. We have developed capabilities for explaining problem-solving actions by incrementally elaborating the control plan underlying the decisions to perform them. We have developed capabilities for acquiring new control knowledge from domain experts automatically.

Our most innovative work on BB1 focused on the idea that reasoning effectively about action requires knowledge about action. In particular, it requires knowledge of: the hierarchy of action types; the patterns of formal parameters defining all actions types; the network of concepts for instantiating formal parameters; the modifiers that can restrict the scopes of defined concept types; the translations of terminal action patterns into executable code; and the partial matches between patterns defining different action types. We developed a body of such knowledge (the ACCORD framework discussed below) for the actions involved in assembling arrangements of objects under constraints. We used the PROTEAN system, which is implemented in BB1, to demonstrate the power of task-specific action knowledge to enhance control, explanation, and learning capabilities. We have also begun to investigate the applicability of this knowledge to another design problem, site layout.

In addition to the two applications mentioned above, several other scientists at Stanford and at other research and industrial laboratories have begun developing application systems in BB1.

[See KSL technical memos KSL 84-16, KSL 85-2, KSL 85-35, KSL 86-38.]

3. *Advanced Architectures:* What kinds of software tools and system architectures can provide orders of magnitude speedup in the performance of expert systems? The Advanced Architectures Project is a long-range project with two related goals:

- To realize a new generation of software system architectures using parallelism to achieve high-speed computation in artificial intelligence applications.

- To specify multiprocessor hardware system architectures that support those parallel computations.

The basic problem we are addressing is to increase the speed of execution of expert systems through the use of parallel computations on a multiprocessor computer system. Part of the effectiveness of expert systems, particularly for real-time applications such as continuous signal data understanding, lies in the speed of execution, or throughput rate. However, for many

significant applications of this type, projected performance limits of uniprocessors fall short of the speed required by as much as several orders of magnitude. Multiprocessor parallel computing must be used to attain the necessary levels of performance.

The anticipated computational requirements for the next decade cannot be realized by just using parallelism at only one particular level of computation (for example, parallel left-hand-side rule matching in rule-based systems). To understand the effectiveness of parallel implementations of expert systems, we must study both the programming problems and the performance issues at all levels of the computational hierarchy:

- The application level.

- The problem-solving framework level.

- The programming language level.

- The hardware system architecture level.

Our research emphasis is therefore on overall software and hardware system architectures for the parallel execution of expert systems.

During the past year, with principal support from DARPA under the Strategic Computing Program, we have demonstrated significant progress at each of the levels. We have also completed the first of a series of "vertical slice" experiments, in which a choice is made at each design level and a simulated execution of the resulting system is analyzed.

Application level

The methodology employed in this project is to select an application and use it as the driver that determines the requirements at the underlying system design levels. That is, the application, or class of applications, should determine the architecture rather than the other way around. Consequently, it is necessary to choose applications very carefully with respect to their complexity, generality and potential for significant speedup.

During the past year we defined and started the development of a new application, within the area of signal understanding, information fusion and situation assessment. The new application, called AIRTRAC, concerns the classification and tracking of light aircraft, some of whom are deliberately trying to evade detection (e.g., smugglers). The sensor data include both acoustic and radar data from distributed sources. Other data include flight plans and intelligence reports. This application has many desirable characteristics, including:

- Multiple sources of input, including both "low-level" (radar, acoustic) and "high-level" (intelligence reports) data;

- Need for both data-driven and model-driven (e.g., using knowledge of intentions or expectations of future behavior) reasoning.

Problem-solving framework level

We have completed the first-pass development of two parallel blackboard framework systems, CAGE and POLIGON. A third framework, CAOS, is complete and has been used for the first vertical slice experiment.

CAGE is an extension of the AGE system that contains concurrency primitives for building parallel constructs. In contrast with POLIGON, CAGE represents a conservative, incremental approach to building parallel systems. CAGE will run both on QLAMBDA and CAOS/CARE simulators, or can be run serially. POLIGON is a demon-driven blackboard system in which all blackboard nodes are active agents. Changes in the blackboard nodes trigger rules to be fired. POLIGON will run on the CAOS/CARE simulator. CAOS ("Concurrent Asynchronous Object System") is a set of language extensions to Lisp for multiprocessor systems. CAOS allows the user to express process and data locality and interprocess communication, organized in an object-oriented manner.

Programming Language Level

We have experimented with Lisp-based languages. One is QLAMBDA (renamed QLISP), an extension to Lisp for a multi-processor, shared memory architecture. (See R. P. Gabriel and J. McCarthy, "Queue-based Multi-processing Lisp," in *Proc. of the 1984 Symposium on Lisp and Functional Programming*, August 1984.) We have also designed and partially implemented a concurrent Lisp for the CARE distributed-memory family of multiprocessor architectures. CAREL (CARE Lisp) is a distributed-memory variant of QLISP. CAREL supports features (like MultipLisp), truly parallel LET binding (like QLISP), active objects with locality and state (like OIL), programmer or automatic specification of locality of computations (like para-functional programming or Flat Concurrent Prolog), and both static assignment of process to processor and dynamic spread of recursive computations through the network via remote function call (like V).

Hardware System Architecture Level

Our activity at this level has been focused on the testing and refinement of CARE, a set of executable system component specifications that can be used to specify a parameterized family of hardware system architectures. The architecture consists of a number of sites interconnected under some specified communication topology. Each site consists of (1) an evaluator (of Lisp forms), (2) an operator that performs message handling, process scheduling, process creation and process synchronization, (3) network ports for message routing, (4) FIFO queues that tie these components together, and (5) a memory -- hence a family of distributed memory machines with a processor of parameterized capability at each node. By specializing nodes to consist only of subsets of the above components, shared memory systems can also be simulated. CARE also permits specification of a suite of instrument probes and display panels that can be used to monitor the simulation of the hardware system.

     E. H. Shortliffe

Vertical Slice Experiment

We completed the first set of comprehensive experiments in implementing and executing knowledge-based expert systems on multiprocessor machines. This experiment consists of running the ELINT application, written in ZetaLisp on CAOS running on the CARE simulator. ELINT is a prototype passive radar signal understanding system that was originally implemented in AGE.

The experiment has two objectives. The first is to investigate the quality of solution and the amount of communication as a function of various degrees of inter-process control. In particular, we are investigating the types and amounts of serialization required to assure an acceptable level of solution quality for ELINT. The second investigates overall execution speedup (or "speeddown") as a function of the number of processors. In particular, a version of ELINT with control adequate to assure satisfactory solution quality was run on simulated CARE arrays ranging in size from four to sixty four processors. The major relations under investigation are execution time and communication behavior as a function of array size. Results of these experiments are currently being analyzed and documented.

[See KSL technical memos KSL-85-24, KSL-86-10, KSL-86-19, KSL-86-20, KSL-86-22, KSL-86-31, KSL-86-41.]

4. *Knowledge Acquisition*: How is knowledge acquired most efficiently from human experts, from observed data, from experience, and from discovery? How can a program discover inconsistencies and incompleteness in its knowledge base? How can the knowledge base be augmented without perturbing the established knowledge base?

Several parallel lines of research on machine learning are in progress, representing a broad spectrum of possibilities for aiding in the construction of new knowledge bases for expert systems. Of these, significant progress was made on two aspects of learning by induction from examples. These two are documented in PhD dissertations by Li-Min Fu and Thomas Dietterich.

Fu's dissertation investigates methods of induction in the context of learning rules and meta-rules for diagnosing cases of jaundice. The program, called RL, uses a rough model, or half-order theory, of the domain in order to guide a systematic search through a space of plausible concept definitions and associations. Experiments show that the quality of rules learned in this fashion are as good as rules derived from texts and physicians through knowledge engineering.

Dietterich's dissertation explores another important problem in theory formation: interpreting observed data in the first place. In the case that an emerging, partially formed theory is used to interpret the data, there is ample opportunity for erroneous extensions to the theory that is being developed to explain the data. We have defined a method for "theory-driven data interpretation" that propagates constraints in order to determine a consistent interpretation of the data. This has been implemented in a program called PRE. The domain in which PRE operates is learning descriptions of UNIX file commands from examples of I/O behavior of a UNIX system in use.

In addition, we have completed a prototype program that serves as a learning

apprentice for systems developed under NEOMYCIN. The model is general, and the program is being tested in NEOMYCIN's domain of medical diagnosis. Its purpose is to "watch" the interaction of an expert diagnosing a difficult case and to build a set of knowledge structures that will allow NEOMYCIN to diagnose similar cases in the same way.

"Chunking" is a learning mechanism that acquires rules from goal-based experience. SOAR is a general problem-solving architecture with a rule-based memory that can use the learning capabilities of chunking for the acquisition and use of macro-operators. Rosenbloom *et al.* are investigating chunking in SOAR and find that chunking obtains extra scope and generality from its intimate connection with the sophisticated problem solver (SOAR) and the memory organization of the production system.

Two MSAI theses (Hewett and Harvey) address learning from human experts. Hewett's program, MARCK, interviews a domain expert to determine why the expert prefers problem-solving actions not chosen by the application system. Harvey's program, WATCH (which is not completely implemented yet), abstracts a domain expert's control heuristics by observing the his or her problem-solving actions. Both of these programs operate in the context of application systems implemented in the BB1 architecture.

[Preliminary results have been published in KSL-85-11, KSL-85-20, KSL-85-26, KSL-85-30, KSL-85-32, KSL-85-34, KSL-85-35, KSL-85-36, KSL-85-38, KSL-85-42, KSL-85-43, KSL-85-44, KSL-85-51, KSL-85-53, KSL-85-56, KSL-86-1, KSL-86-6, KSL-86-7, KSL-86-35, KSL-86-38.]

5. *Knowledge Utilization*: By what inference methods can many sources of knowledge of diverse types be made to contribute jointly and efficiently toward solutions? How can knowledge be used intelligently, especially in systems with large knowledge bases, so that it is applied in an appropriate manner at the appropriate time?

A PhD dissertation by Greg Cooper has been completed in which a model of inexact reasoning is proposed and demonstrated using both probabilistic and causal knowledge. The key idea is that estimates of probability ranges can be modified by using knowledge of causal relations.

[See KSL technical memos KSL-85-14, KSL-85-18, KSL-85-23, KSL-85-25, KSL-85-35, KSL-85-40, KSL-85-41, KSL-85-46, KSL-86-26, KSL-86-30.]

6. *Software Tools*: How can specific programs that solve specific problems be generalized to more widely useful tools to aid in the development of other programs of the same class?

We have continued the development of new software tools for expert system construction and the distribution of packages that are reliable enough and documented so that other laboratories can use them. These include the old rule-based EMYCIN system, MRS, and AGE.

We have continued our development and refinement of BB1, including the following new capabilities: generic control knowledge sources for refining abstract control plans; graphical display of the dynamic control plan and other system-state information; strategic explanation of problem-solving actions; two programs for automatically learning control heuristics from experts; general knowledge-base facilities; and "mouse-controlled" interfaces for all system functions.

E. H. Shortliffe

We also have extended BB1 to exploit any user-defined "framework" embodying task-specific action knowledge (see "Progress" above). We have implemented the ACCORD framework for systems that assemble arrangements of objects under constraints. We have released BB1 to approximately fifteen research groups outside of Stanford.

The ACCORD language allows definition of knowledge sources at a high level of description and represents a significant improvement in clarity and ease of definition of knowledge sources. We plan to release ACCORD sometime during the summer of 1986.

[See KSL technical memos KSL-85-12, KSL-85-15, KSL-86-38.]

7. *Explanation and Tutoring*:   How can the knowledge base and the line of reasoning used in solving a particular problem be explained to users? What constitutes a sufficient or an acceptable explanation for different classes of users? How can knowledge in a system be transferred effectively to students and trainees?

We have been concerned for years about the understandability of expert systems. We are currently focusing on the high resolution bit-mapped displays on Lisp workstations as a desirable mode of explaining the contents of knowledge bases. A prototype program, called GUIDON-WATCH, has been written and documented that provides easily understood windows into the problem solving activities of NEOMYCIN.

A knowledge-based system must not only be able to recommend an action but also must provide an explanation as to why that action is the most desirable and what task or subtask the proposed action will accomplish. Control knowledge in BB1 has a hierarchical structure of *heuristics*, a current *focus*, and one or more levels of *strategy* for solving the problem. Since BB1's control knowledge is explicitly represented as knowledge sources, an explanation of the problem solving can be constructed at each level of abstraction. BB1 offers several ways for the user to determine the rationale behind its recommendations, including the "Explain", "Describe", and "Why" commands.

Jeff Harvey designed and built the "Why" facility to run under BB1. It differs from those found in rule based systems because it explains why it recommends an action, rather than just explaining why the system is asking a question. When first asked "Why", the system describes the heuristics used to evaluate feasible actions. On additional "Why" queries, the system describes control decisions in more generality, continuing until the highest level of abstraction is reached. The "Describe" command is similar to "Why", but all of the control decisions are explained at the present level of abstraction.

"Explain" gives the rationale for why the decision is a good thing to do, in terms of the ratings of that action by each of the active heuristics. The action with the highest rating is the one recommended by BB1.

[See KSL technical memos KSL-85-39, KSL-86-2, KSL-86-15, KSL-86-34.]

8. *Planning and Design*:   What are reasonable and effective methods for planning and design?   How can symbolic knowledge be coupled with numerical constraints? How are constraints propagated in design problems?

We have made significant progress in this last year in merging AI and decision-analytic methods in developing plans. This is largely reflected in the program named ONYX, which "backs up" the reasoning in ONCOCIN with planning at a more fundamental level.

[See KSL technical memos KSL-85-10, KSL-85-52, KSL-85-55.]

9. *Diagnosis and Therapy Management*: How can we build a diagnostic system that reflects any of several diagnostic strategies? How can we use knowledge at different levels of abstraction in the diagnostic process? How can a rational therapy plan be devised that is tailored to the specifics of an individual case?

ONCOCIN (see separate section on ONCOCIN) is the primary vehicle for studying therapy management, and substantial progress has been made in representing and using therapy plans.

[See KSL technical reports: KSL-85-21, KSL-85-22, KSL-85-29, KSL-85-33, KSL-85-50, KSL-86-3, KSL-86-4, KSL-86-5, KSL-86-9.]

Relevant Publications

KSL 85-10   Curtis Langlotz, Lawrence Fagan, Samson Tu, John Williams, and Branimir Sikic; **ONYX: An Architecture for Planning in Uncertain Environments,** May 1985. **7 pages**

KSL 85-11   Shoko Tsuji, Edward H. Shortliffe; **Graphics for Knowledge Engineers: A Window on Knowledge Base Management,** April 1985. **23 pages**

KSL 85-12   Stuart J. Russell; **The Compleat Guide to MRS,** June 1985. **121 pages**

KSL 85-14   Matthew L. Ginsberg; **Implementing Probabilistic Reasoning,** April 1985. **12 pages**

KSL 85-15   Lane, Differding, and Shortliffe; **Graphical Access to Medical Expert Systems: II. Design of an Interface for Physicians,** July 1985. **22 pages**

KSL 85-16   (Working Paper) William J. Clancey; **Representing Control Knowledge as Abstract Tasks and Metarules,** April 1985. **56 pages**

KSL 85-17   Mark Musen, Curtis Langlotz, Lawrence Fagan, and Edward Shortliffe; **Rationale for Knowledge Base Redesign in a Medical Advice System,** April 1985. **6 pages**

KSL 85-18   Vineet Singh, and Michael Genesereth; **PM: A Parallel Execution Model for Backward-Chaining Deductions,** August 1985. Submitted for publication to: *Future Computing Systems,* 1985. **26 pages**

KSL 85-19   Mark H. Richer; **An Evaluation of Knowledge-Based Software Tools,** May 1985. **21 pages**

KSL 85-20   STAN-CS-85-1068. Mark H. Richer and William J. Clancey; **GUIDON-WATCH: A Graphic Interface for Browsing and Viewing a Knowledge Based System,** September 1985. *IEEE Computer Graphics and Applications,* Vol. 4, No. 11, November 1985, pp. 51-64. **33 pages**

KSL 85-21   D.H. Hickam, E.H. Shortliffe, M.B. Bischoff, A.C. Scott, C.D. Jacobs; **A Study of the Treatment Advice of a Computer-Based Cancer Chemotherapy Protocol Advisor,** July 1985. To appear in the: *Annals of Internal Medicine,* 1985. **34 pages**

E. H. Shortliffe

KSL 85-22 D.L. Kent, E.H. Shortliffe, R.W. Carlson, M.B. Bischoff, C.D. Jacobs; **Improvements in Data Collection Through Physician Use of a Computer-Based Chemotherapy Treatment Consultant**, May 1985. **20 pages**

KSL 85-23 (Working Paper) David Heckerman; **Probabilistic Interpretations for MYCIN's Certainty Factors**, May 1985. **28 pages**

KSL 85-24 Penny Nii; **Research on Blackboard Architectures at the Heuristic Programming Project**, May 1985. **11 pages**

KSL 85-25 Matthew L. Ginsberg; **Decision Procedures**, September 1985. **22 pages**

KSL 85-26 David C. Wilkins, William J. Clancey, & Bruce G. Buchanan; **An Overview of the Odysseus Learning Apprentice**, August 1985. *Machine Learning: A Guide to Current Research, Academic Press*, 1986. **4 pages**

KSL 85-27 (Working Paper)] G.D. Rennels, E.H. Shortliffe; **Medical Advice Systems.** May 1985. To appear in: *Encyclopedia of Artificial Intelligence,* published by John Wiley & Sons. **16 pages**

KSL 85-28 (Working Paper) E.H. Shortliffe; **The State of the Art of the Science.** *Proceedings of the Conference on Medical Information Sciences,* U. of Texas Health Science Center at San Antonio, July 1985. **10 pages**

KSL 85-29 Michael G. Walker, Robert Blum, and Lawrence M. Fagan; **MINIMYCIN: A Miniature Rule-Based System.** Published in: *Clinical Computing,* Vol. 2, No. 4, 1985. **8 pages**

KSL 85-30 Differing, Combs, Musen, Lane, Fagan, & Shortliffe; **Graphical Access to Medical Expert Systems: III Design of a Knowledge Aquisition Environment**, August 1985.

KSL 85-31 William J. Clancey; **Review of Sowa's "Conceptual Structures"**, August 1985. To appear in the: *Journal of Artificial Intelligence,* 1985. **12 pages**

KSL 85-32 (Working Paper) Timothy F. Thompson and William J. Clancey; **The CASTER System: An Experiment in Knowledge Acquisition Within a Generic Expert System Shell**, August 1985. **25 pages**

KSL 85-33 James F. Brinkley; **Knowledge Driven Ultrasonic Three-Dimensional Organ Modelling**, August 1985. Published in: *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. PAMI-7, No. 4, pp 431-441. July 1985. **11 pages**

KSL 85-34 John Laird, Paul Rosenbloom, & Allen Newell; **Chunking in SOAR: The Anatomy of a General Learning Mechanism**, September 1985. In: *Machine Learning,* Vol. 1, No. 1, 1986. Also appears as Xerox PARC ISL-13 and CMU-CS-85-154. **34 pages**

KSL 85-35 Barbara Hayes-Roth, Bruce Buchanan, Olivier Lichtarge, Mike Hewett, Russ Altman, James Brinkley, Craig Cornelius, Bruce Duncan, and Oleg Jardetzky; **Elucidating Protein Structure from Constraints in PROTEAN**, October 1985. Submitted for publication to: *Insight Series in Applied AI.* **29 pages**

KSL 85-36 (Working Paper) Peter Karp; **Thesis Proposal: Qualitative Simulation and Discovery in Molecular Biology**, September 1985. **31 pages**

KSL 85-37 Bruce G. Buchanan; **EXPERT SYSTEMS: Working Systems and the Research Literature**, December 1985. Appears in: *Expert Systems.* **55 pages**

KSL 85-38   Bruce G. Buchanan; **Some Approaches to Knowledge Acquisition**, October 1985. Appears in: *Proceedings of the Third Int. Workshop on Machine Learning.* **5 pages**

KSL 85-39   (Working Paper) Glenn D. Rennels, Edward H. Shortliffe and Perry L. Miller; **A Model of Choice and Explanation in Medical Management**, October 1985. Appears in: *Computers and Biomedical Research.* **17 pages**

KSL 85-40   (Thesis) Jeffrey S. Rosenschein; **Rational Interaction: Cooperation Among Intelligent Agents**, October 1985. **133 pages**

KSL 85-41   Buchanan, B.G., Hayes-Roth, B., Lichtarge, O., Altman, A., Brinkley, J., Hewitt, M., Cornelius, C., Duncan, B., and Jardetzky, O.: **The Heuristic Refinement Method for Deriving Solution Structures of Proteins**, March 1986. In: *Proceedings of the National Academy of Science.* **10 pages**

KSL 85-42   Li-Min Fu and Bruce G. Buchanan; **Inductive Knowledge Acquisition for Rule-Based Expert Systems**, October 1985. Submitted for publication: *Artificial Intelligence.* **34 pages**

KSL 85-43   (Working Paper) Robert L. Blum; **Two Stage Regression: Application to a Time-Oriented Clinical Database**, October 1985. To appear in: *Statistics in Medicine.* **27 pages**

KSL 85-44   (Thesis) Li-Min Fu; **Learning Object-Level and Meta-Level Knowledge in Expert Systems**, November 1985. **229 pages**

KSL 85-46   (Working Paper) Richard Treitel and Michael R. Genesereth; **Choosing Directions for Rules**, March 1986. Accepted for publication by *AAAI-86.* Submitted for publication to: *Journal of Automated Reasoning.* **36 pages**

KSL 85-50   (Journal Memo) G.D. Rennels, E.H. Shortliffe, F.E. Stockdale and P.L. Miller; **Reasoning from the Clinical Literature: a "Distance" Metric,** November 1985. To appear in: *Proceedings of AAMSI Congress 86,* Anaheim, CA, May 1986. **6 pages**

KSL 85-51   (Journal Memo) M.A. Musen, J.A. Rohn, L.M. Fagan and E.H. Shortliffe; **Knowledge Engineering for a Clinical Trial Advice System: Uncovering Errors in Protocol Specification**, November 1985. To appear in: *Proceedings of AAMSI Congress 86,* Anaheim, CA, May 7-10, 1985. **5 pages**

KSL 85-52   (Journal Memo) C.P. Langlotz, L.M. Fagan and E.H. Shortliffe; **Overcoming Limitations of Artificial Intelligence Planning Techniques**, November 1985. To appear in: *Proceedings of AAMSI Congress 86,* Anaheim, CA, May 8-10, 1986. **5 pages**

KSL 85-53   (Journal Memo) M.A. Musen, L.M. Fagan and E.H. Shortliffe; **Graphical Specification of Procedural Knowledge for an Expert System**, December 1985. To be presented at: *Second IEEE Computer Society Workshop on Visual Languages,* Dallas, TX, June 1986. **18 pages**

KSL 85-54   Devika Subramanian and Bruce G. Buchanan; **A General Reading List for Artificial Intelligence**, December 1985. **63 pages**

KSL 85-55   (Working Paper) C.P. Langlotz, L.M. Fagan, S.W. Tu and B.I. Sikic; **An Architecture for Planning under Uncertainty**, December 1985. **19 pages**

KSL 85-56   (Journal Memo) D.M. Combs, M.A. Musen, L.M. Fagan and E.H. Shortliffe; **Graphical Entry of Procedural and Inferential Knowledge,**

December 1985. To appear in: *Proceedings of AAMSI Congress 1986* 5 **pages**

KSL 86-1   (Journal Memo) M.A. Musen, L.M. Fagan, D.M. Combs and E.H. Shortliffe; **Facilitating Knowledge Entry for an Oncology Therapy Advisor Using a Model of the Application Area,** January 1986. Accepted for publication in: *Proceedings, MEDINFO 86.* **9 pages**

KSL 86-2   (Journal Memo) E. Horvitz, D. Heckerman, B. Nathwani and L. Fagan; **The Use of a Heuristic Problem-Solving Hierarchy to Facilitate the Explanation of Hypothesis-Directed Reasoning,** January 1986. Accepted for publication in: *Proceedings, MEDINFO 86.* **5 pages**

KSL 86-3   (Journal Memo) C.P. Langlotz, L.M. Fagan, S.W. Tu, B.I. Sikic and E.H. Shortliffe; **Combining Artificial Intelligence and Decision Analysis for Automated Therapy Planning Assistance,** January 1986. Accepted for publication in: *Proceedings, MEDINFO 86.* **5 pages**

KSL 86-4   (Journal Memo) M.G. Kahn, L.M. Fagan and E.H. Shortliffe; **Context-Specific Interpretation of Patient Records for a Therapy Advice System,** January 1986. Accepted for publication in: *Proceedings, MEDINFO 86.* **5 pages**

KSL 86-5   (Journal Memo) G.D. Rennels, E.H. Shortliffe, F.E. Stockdale and P.L. Miller; **Reasoning from the Clinical Literature: The Roundsman System,** January 1986. Accepted for publication in: *Proceedings, MEDINFO 86.* **5 pages**

KSL 86-6   (Journal Memo) S.M. Downs, M.G. Walker and R.L. Blum; **Automated Summarization of On-Line Medical Records,** January 1986. Accepted for publication in: *Proceedings, MEDINFO 86.* **5 pages**

KSL 86-7   (Journal Memo) M.G. Walker and R.L. Blum; **Towards Automated Discovery from Clinical Databases: the RADIX Project,** January 1986. Accepted for publication in: *MEDINFO 86* **5 pages**

KSL 86-9   (Journal Memo) Gregory F. Cooper; **A Diagnostic Method That Uses Casual Knowledge and Linear Programming in the Application of Bayes' Formula,** January 1986. Accepted for publication in: *Computer Methods and Programs in Biomedicine.* **24 pages**

KSL 86-10   (Working Paper) James Rice; **The Poligon User's Manual,** February 1986. **68 pages**

KSL 86-11   (Working Paper) William J. Clancey; **From Guidon to Neomycin and Heracles in Twenty Short Lessons,** February 1986. **48 pages**

KSL 86-15   (Working Paper) William J. Clancey; **Qualitative Student Models,** February 1986. Submitted for publication to: *First Annual Review of Computer Science.* **88 pages**

KSL 86-17   Robert S. Engelmore & Craig W. Cornelius; **Heuristic Programming Project, October 1982 - September 1985, Final Report,** February 1986. **16 pages**

KSL 86-19   J. P. Rice; **Poligon, A System for Parallel Problem Solving,** April 1986. To appear in Proceedings of DARPA Workshop on Expert Systems Technology Base, Asilomar, April 1986. **19 pages**

KSL 86-20   J.R. Delaney; **Multi-System Report Integration Using Blackboards,** March 1986. Accepted for publication in: *1986 American Control Conference.* **12 pages**

KSL 86-22   Eric Schoen; The CAOS System, March 1986. To appear in Proceedings of DARPA Workshop on Expert Systems Technology Base, Asilomar, April 1986. 70 pages

KSL 86-26   (Working Paper) C.P. Langlotz, E.H. Shortliffe and L.M. Fagan; Using Decision Theory to Justify Heuristics, March 1986. Accepted for publication in: *AAAI-86* 15 pages

KSL 86-27   (Working Paper) William J. Clancey; The Science and Engineering of Qualitative Models, March 1986. Submitted for publication to: *AAAI-86, Science Track: Applications, Philosophical and Scientific Foundations.* 20 pages

KSL 86-30   (Working Paper) David C. Wilkins & Bruce G. Buchanan; On Optimizing Rule Sets When Reasoning Under Uncertainty, April 1986. Submitted for publication to: *AAAI-86*. 14 pages

KSL 86-31   Nelleke Aiello; User-Directed Control of Parallelism: The CAGE System, April 1986. To appear in Proceedings of DARPA Workshop on Expert Systems Technology Base, Asilomar, April 1986. 12 pages

KSL 86-32   (Working Paper) Peter D. Karp & David C. Wilkins; An Analysis of the Deep/Shallow Distinction for Expert Systems, April 1986. 18 pages

KSL 86-34   (Working Paper) William J. Clancey, Mark Richer, David C. Wilkins, Steve Barnhouse, Curt Kapsner, David Leserman, John Macias, Arif Merchant and Naomi Rodolitz; Guidon-Debug: The student as knowledge engineer, April 1986. 17 pages

KSL 86-35   (Working Paper) Michael G. Walker; How Feasible is Automated Discovery? April 1986. Submitted for publication to: *IEEE Expert.* 24 pages

KSL 86-41   H. Penny Nii; CAGE and POLIGON: Two Frameworks for Blackboard-based Concurrent Problem Solving, April 1986. To appear in Proceedings of DARPA Workshop on Expert Systems Technology Base, Asilomar, April 1986. 9 pages

Funding Support

Project Title: Heuristic Programming Project
Principal Investigators: Edward A. Feigenbaum and Bruce G. Buchanan
Amount: $3,354,493
Period Covered: 10/1/82 - 9/30/85 (Note: New three-year contract in negotiation.)

Agency: Defense Advanced Research Projects Agency; MDA903-83-C-0188
Project Title: Research Computing Equipment Modernization
Principal Investigator: Edward A. Feigenbaum
Amount: $2,565,000
Period Covered: 6/1/83 - 5/31/86

Agency: Defense Advanced Research Projects Agency; F30602-85-C-0012
Project Title: Expert Systems on Multiprocessor Architecture
Principal Investigator: Edward A. Feigenbaum
Award Amount: $1,873,511
Period Covered: 3/14/85 - 3/13/87

Agency: Lawrence Livermore
Project Title: Research on Intelligent Budget Planning and Resource Management Systems
Principal Investigator: Bruce G. Buchanan
Award Amount: 124,905
Period Covered: 12/14/84 - 9/30/86

Agency: Josiah Macy, Jr. Foundation
Project Title: A Family of Intelligent Tutoring Programs for Medical Diagnosis
Principal Investigator: Bruce G. Buchanan
Award Amount: $503,415
Period Covered: 3/1/85 - 2/29/88

Agency: Martin-Marietta Corporation
Project Title: Intelligent Task Automation
Principal Investigator: Michael Genesereth
Period Covered: 1/1/85 - 12/31/85

Agency: NASA-Ames Research Center
Project Title: Research On Knowledge Representation
Principal Investigator: Bruce G. Buchanan
Amount: $343,144
Period Covered: 10/1/83 - 12/31/87

Agency: NASA-AMES Research Center; NCC 2-220, S1
Project Title: Research on Advanced Knowledge-based System Architectures
Principal Investigator: Edward A. Feigenbaum
Award Amount: $381,417
Period Covered: 10/1/82 - 1/31/87

Agency: National Science Foundation; MCS-8310236
Project Title: Applications of AI to Molecular Biology
Principal Investigator: Edward A. Feigenbaum
Award Amount: $405,836
Period Covered: 11/1/83 - 10/31/86

Agency: National Science Foundation

Project Title: Information Structure-Use Knowledge-Based Expert Systems
Principal Investigator: Bruce Buchanan & Ted Shortliffe
Award Amount: $333,710
Period Covered: 3/15/84 - 2/28/87

Agency: National Science Foundation
Project Title: Interpretation of NMR Data for Proteins Using AI
Methods
Principal Investigator: Bruce Buchanan & Oleg Jardetzky
Award Amount: $100,000
Period Covered: 11/1/84 - 10/31/86

Agency: National Science Foundation; MCS-8303142
Project Title: The Mechanization of Formal Reasoning (Computer Research)
Principal Investigator: Edward A. Feigenbaum
Award Amount: $286,144
Period Covered: 7/15/83 - 6/30/86

Agency: Office of Naval Research
Project Title: Computer-Based Tutors for Explaining and Managing
the Process of Diagnostic Reasoning
Principal Investigator: Bruce G. Buchanan
Award Amount: $513,662
Period Covered: 3/15/85 - 3/14/88

Agency: Office of Naval Research [Pending]
Project Title: Expert Control of Problem-Solving Search
Principal Investigator: Bruce G. Buchanan
Award Amount: $725,899
Period Covered: 8/1/85 - 7/31/88

Agency: Rockwell International [Pending]
Project Title: Intelligent Real-Time Process Control
Principal Investigators: Edward A. Feigenbaum
Amount: $643,059
Period Covered: 1/1/87-12/31/89

Agency: Boeing Computer Services Company [Pending]
Project Title: Research on Blackboard Systems and Intelligent
Resource Management Systems
Principal Investigators: E.A. Feigenbaum and B.G. Buchanan
Amount Requested: $569,604
Period Covered: 4/1/86-3/31/88

Agency: Office of Naval Research [Pending]
Project Title: Modeling Expert Control Knowledge
Principal Investigator: Bruce G. Buchanan
Amount Requested: $96,909
Period Covered: 11/1/85 - 10/31/86
Percentage Effort Committed to Project: 3%, no salary support

Agency: Defense Advanced Research Projects Agency [Pending]
Project Title: Research on Intelligent Assistants for Protein Chemists
Principal Investigator: Bruce G. Buchanan, Oleg Jardetzky, Hans
Andersen, Harden McConnell
Amount Requested: $7,962,678

Period Covered: 10/1/86-9/30/91

Agency: National Science Foundation [Pending]
Project Title: Heuristic Refinement for Deriving Solution
Structures of Proteins
Principal Investigator: Bruce G. Buchanan and Oleg Jardetzky
Amount Requested: $589,764
Period Covered: 11/1/86-10/31/89

Interactions with the SUMEX-AIM Resource

Our interactions with the SUMEX-AIM resource involve the facilities -- both hardware
and software -- and the staff -- both technical and administrative. Taken together as a
whole resource, they constitute an essential part of the research structure for the KSL.
Many of the grants and contracts from other agencies have been awarded partly because
of the cost-effectiveness of AI research in the KSL due to the fact that much of our
computing needs could be more than adequately met by the SUMEX-AIM resource.

We rely on the central SUMEX facility as a focal point for all the research within the
KSL, not only for much of our computing, but for communications and links to our
many collaborators as well. As a common communications medium alone, it has
significantly enhanced the nature of our work and the reach of our collaborations. The
existence of the central time-shared facility has allowed us to explore new ideas at very
small incremental cost.

As SUMEX and the KSL acquire a diversity of hardware, including LISP workstations
machines and smaller personal computers, we rely more and more heavily on the
SUMEX staff for integration of these new resources into the local network system. The
staff has been extremely helpful and effective in dealing with the myriad of complex
technical issues and leading us competently into this world of decentralized, diversified
computing. At the same time, the staff has provided a stable, efficient central time-
shared machine running software that has been developed at many sites over many
years. Without the dedication of the SUMEX staff, the KSL would not be at the
forefront of AI research.

III.A.3.5. Training Activities

The SUMEX resource exists to facilitate biomedical artificial intelligence applications from program development through testing in the target research communities. This user orientation on the part of the facility and staff has been a unique feature of our resource and is responsible in large part for our success in community building. The resource staff has spent significant effort in assisting users gain access to the system and use it effectively. We have also spent substantial effort to develop, maintain, and facilitate access to documentation and interactive help facilities. The HELP and Bulletin Board subsystems have been important in this effort to help users get familiar with the computing environment.

On another front, we have regularly accepted a number of scientific visitors for periods of several months to a year, to work with us to learn the techniques of expert system definition and building and to collaborate with us on specific projects. Our ability to accommodate such visitors is severely limited by space, computing, and manpower resources to support such visitors within the demands of our on-going research.

Finally, the training of graduate students is an essential part of the research and educational activities of the KSL. Currently, 41 students are working with our projects centered in Computer Science and another 20 students are working with the Medical Computer Science program in Medicine. Of the 41 working in Computer Science, 25 are working toward Ph.D. degrees, and 16 are working toward M.S. degrees. A number of students are pursuing interdisciplinary programs and come from the Departments of Engineering, Mathematics, Education, and Medicine.

Based on the SUMEX-AIM community environment, we have initiated two unique and special academic degree programs at Stanford, the Medical Information Science program and the Masters of Science in AI, to increase the number of students we produce for research and industry, who are knowledgeable about knowledge-based system techniques.

The *Medical Information Sciences (MIS)* program is one of the most obvious signs of the local academic impact of the SUMEX-AIM resource. The MIS program received recent University approval (in October 1982) as an innovative training program that offers MS and PhD degrees to individuals with a career commitment to applying computers and decision sciences in the field of medicine. The MIS training program is based in the School of Medicine, directed by Dr. Shortliffe, co-directed by Dr. Fagan, and overseen by a group of nine University faculty that includes several faculty from the Knowledge Systems Laboratory (Profs. Shortliffe, Feigenbaum, Buchanan, and Genesereth). It was Stanford's active ongoing research in medical computer science, plus a world-wide reputation for the excellence and rigor of those research efforts, that persuaded the University that the field warranted a new academic degree program in the area. A group of faculty from the medical school and the computer science department argued that research in medical computing has historically been constrained by a lack of talented individuals who have a solid footing in both the medical and computer science fields. The specialized curriculum offered by the new program is intended to overcome the limitations of previous training options. It focuses on the development of a new generation of researchers with a commitment to developing new knowledge about optimal methods for developing practical computer-based solutions to biomedical needs.

The program accepted its first class of four trainees in the summer of 1983 and a second class of five entered last summer. A third group of seven students has just been selected to begin during 1985. The proposed steady state size for the program (which should be reached in 1986) is 20-22 trainees. Applicants to the program in our first two years have come from a number of backgrounds (including seven MD's and five medical students). We do not wish to provide too narrow a definition of what kinds of

prior training are pertinent because of the interdisciplinary nature of the field. The program has accordingly encouraged applications from any of the following:

- medical students who wish to combine MD training with formal degree work and research experience in MIS;

- physicians who wish to obtain formal MIS training after their MD or their residency, perhaps in conjunction with a clinical fellowship at Stanford Medical Center;

- recent BA or BS graduates who have decided on a career applying computer science in the medical world;·

- current Stanford undergraduates who wish to extend their Stanford training an extra year in order to obtain a "co-terminus" MS in the MIS program;

- recent PhD graduates who wish post-doctoral training, perhaps with the formal MS credential, to complement their primary field of training.

In addition, a special one-year MS program is available for established academic medical researchers who may wish to augment their computing and statistical skills during a sabbatical break.

With the exception of this latter group, all students spend a minimum of two years at Stanford (four years for PhD students) and are expected to undertake significant research projects for either degree. Research opportunities abound, however, and they of course include the several Stanford AIM projects as well as research in psychological and formal statistical approaches to medical decision making, applied instrumentation, large medical databases, and a variety of other applications projects at the medical center and on the main campus. Several students are already contributing in major ways to the AIM projects and core research described in this application.

Early evidence suggests that the program already has an excellent reputation due to:

- high quality students, many of whom are beginning to publish their work in conference proceedings and refereed journals;

- a rigorous curriculum that includes newly-developed course offerings that are available to the University's medical students, undergraduates, and computer science students as well as to the program's trainees;

- excellent computing facilities combined with ample and diverse opportunities for medical computer science and medical decision science research;

- the program's great potential for a beneficial impact upon health care delivery in the highly technologic but cost-sensitive era that lies ahead.

The program has been successful in raising financial and equipment support (almost $1M in hardware gifts from Hewlett Packard, Xerox, and Texas Instruments; over $200K in cash donations from corporations and foundations; and an NIH post-doctoral training grant from the National Library of Medicine).

The *Master of Science in Computer Science: Artificial Intelligence (MS:AI)* program is a terminal professional degree offered for students who wish to develop a competence in the design of substantial knowledge-based AI applications but who do not intend to obtain a Ph.D. degree.   The MS:AI program is administered by the Committee for Applied Artificial Intelligence, composed of faculty and research staff of the Computer Science Department.   Normally, students spend two years in the program with their

time divided equally between course work and research. In the first year, the emphasis
is on acquiring fundamental concepts and tools through course work and and project
involvement. During the second year, students implement and document a substantial
AI application project.

III.A.3.6. Resource Operations and Usage

The following data give an overview of various aspects of SUMEX-AIM resource usage. There are 5 subsections containing data respectively for:

1. Overall resource loading data (page 57).

2. Relative system loading by community (page 58).

3. Individual project and community usage (page 61).

4. Network usage data (page 69).

5. System reliability data (page 69).

For the most part, the data used for these plots cover the entire span of the SUMEX-AIM project. This includes data from both the KI-TENEX system and the current DECsystem 2060. At the point where the SUMEX-AIM community switched over to the 2060 (February, 1983), you will notice sharp changes in most of the graphs. This is due to many reasons briefly mentioned here:

1. Even though the TENEX operating system used on the KI-10 was a forerunner of the current Tops20 operating system, the Tops20 system is still different from TENEX is many ways. Tops20 uses a radically different job scheduling mechanism, different methods for computing monitor statistics, different I/O routines, etc. In general, it can not be assumed that statistics measured on the TENEX system correlate one to one with similar statistics under Tops20.

2. The KL-10 processor on the 2060 is a faster processor than the KI-10 processor used previously. Hence, a job running on the KL-10 will use less CPU time than the same job running on the KI-10. This aspect is further complicated by the fact that the SUMEX KI-10 system was a dual processor system.

3. The SUMEX-AIM Community was changing during the time of the transfer to the 2060. The usage of the GENET community on SUMEX had just been phased out. This part of the community accounted for much of the CPU time used by the AIM community. Since the purchase of the 2060 was partially funded by the Heuristic Programming Project (HPP), an additional number of HPP Core Research Projects started using the 2060, increasing the Stanford communities usage of the machine.

4. Finally, the move to the 2060 occurred during a pivotal time in the community when more and more projects were either moving to their own local timesharing machines, or onto specialized Lisp workstations. It also was the time for the closure of many long time SUMEX-AIM projects, like DENDRAL and PUFF/VM.

Any conclusions reached by comparing the data before and after February, 1983 should be done with caution. The data is included in this year's annual report mostly for casual comparison.

Overall Resource Loading Data

The following plot displays total CPU time delivered per month. This data includes usage of the KI-TENEX system and the current DECsystem 2060.
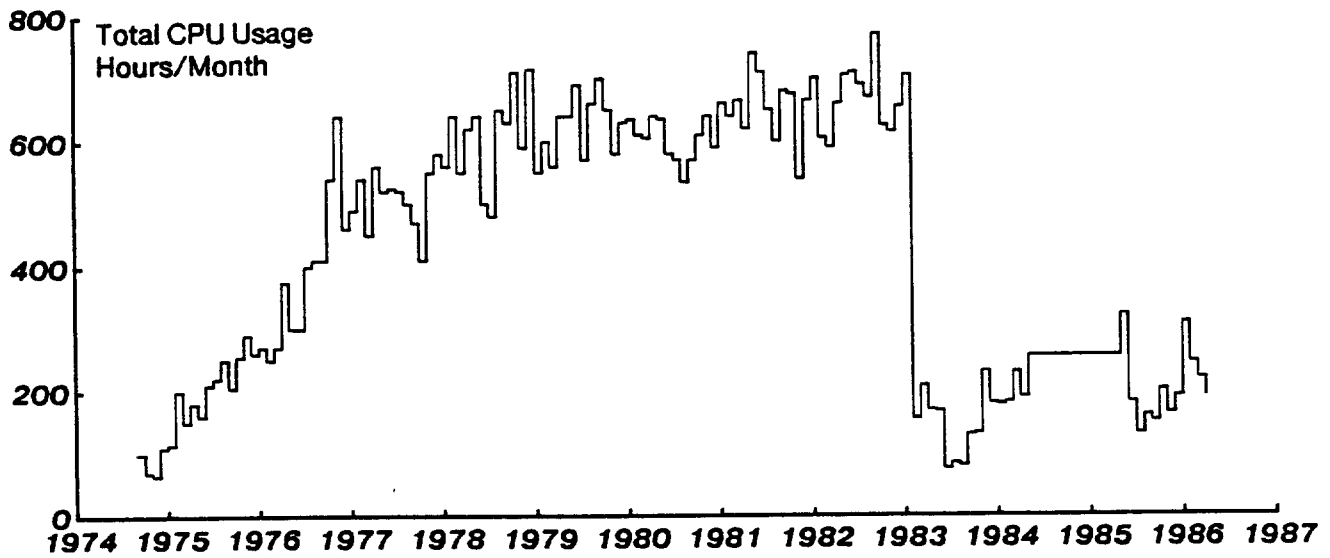
Figure 7: Total CPU Time Consumed by Month

E. H. Shortliffe

Relative System Loading by Community

The SUMEX resource is divided, for administrative purposes, into three major communities: user projects based at the Stanford Medical School (*Stanford Projects*), user projects based outside of Stanford (*National AIM Projects*), and common system development efforts (*System Staff*). As defined in the resource management plan approved by the BRP at the start of the project, the available system CPU capacity and file space resources are divided between these communities as follows:

Stanford	40%
AIM	40%
Staff	20%

The "available" resources to be divided up in this way are those remaining after various monitor and community-wide functions are accounted for. These include such things as job scheduling, overhead, network service, file space for subsystems, documentation, etc.

The monthly usage of CPU resources and terminal connect time for each of these three communities relative to their respective aliquots is shown in the plots in Figure 8 and Figure 9. As mentioned on page 56, these plots include both KI-10 and 2060 usage data.
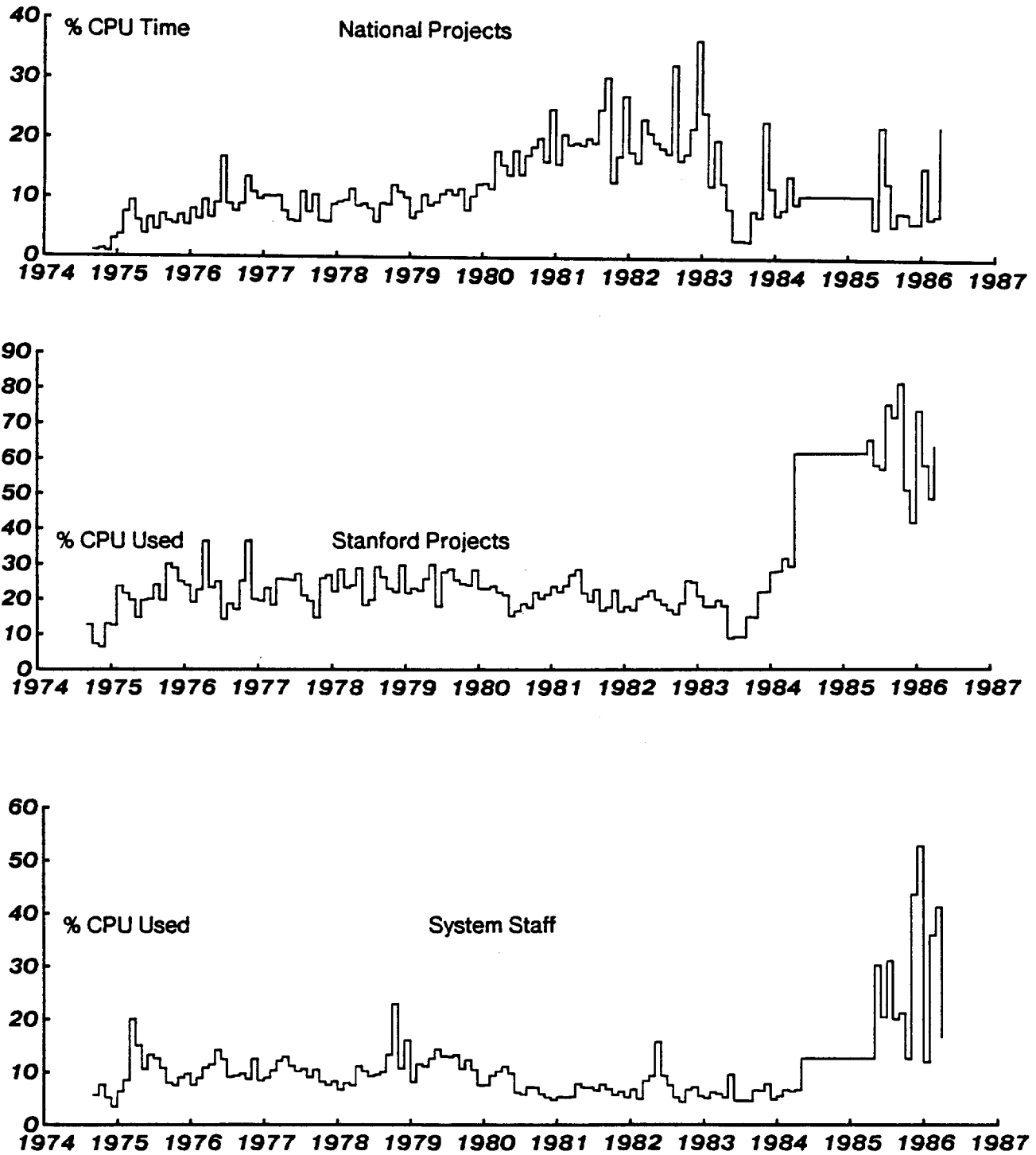
Figure 8:   Monthly CPU Usage by Community
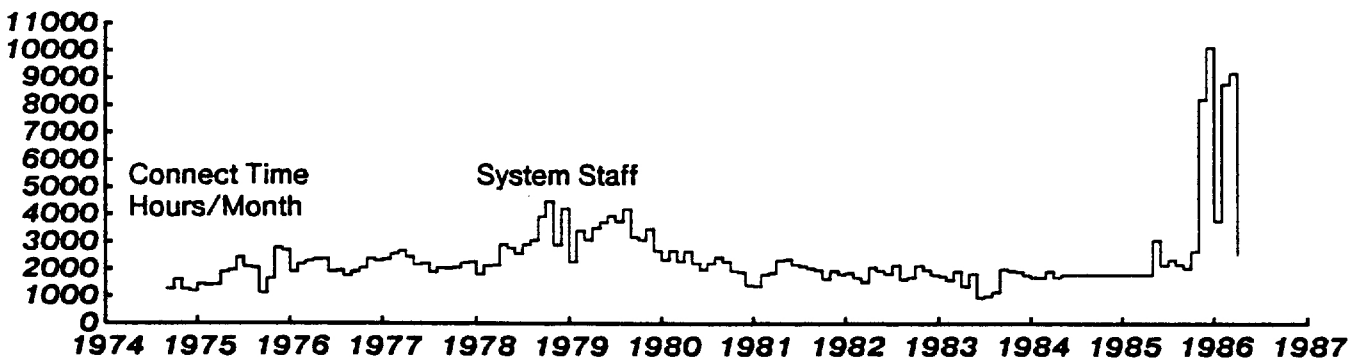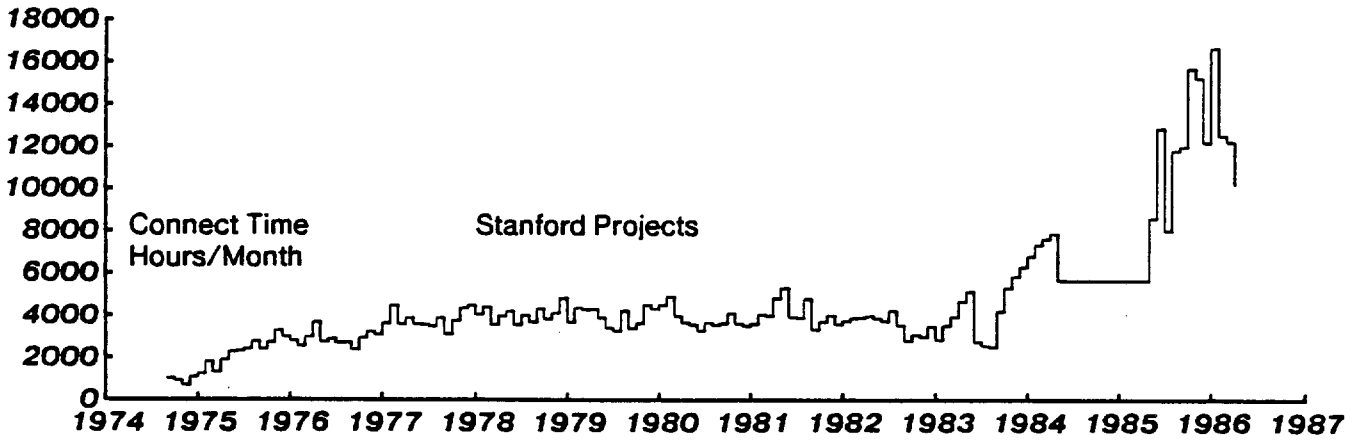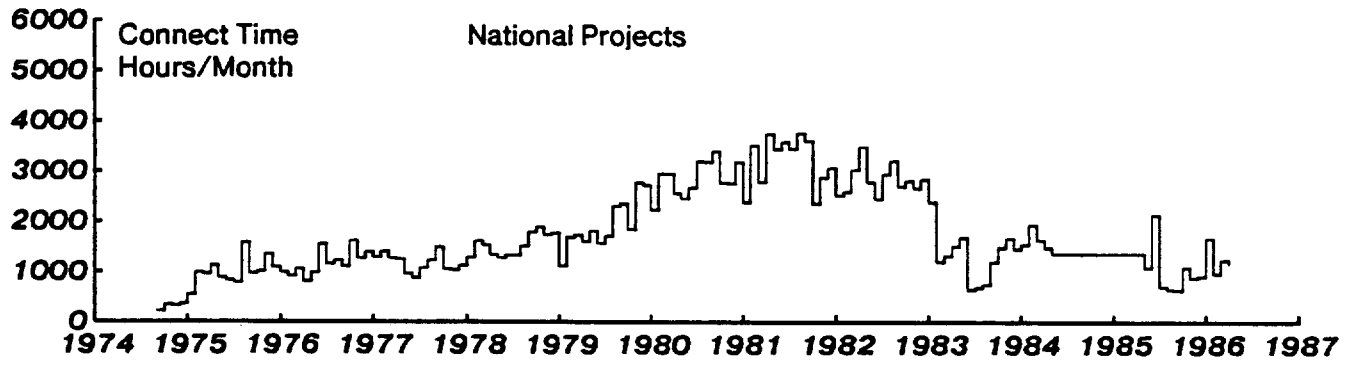
Figure 9:   Monthly Terminal Connect Time by Community

Individual Project and Community Usage

The following histogram and table show cumulative resource usage by collaborative project and community during the past grant year. The histogram displays the project distribution of the total CPU time consumed between May 1, 1985 and April 30, 1986, on the SUMEX-AIM DECsystem2060 system.

In the table following, entries include a text summary of the funding sources (outside of SUMEX-supplied computing resources) for currently active projects, total CPU consumption by project (Hours), total terminal connect time by project (Hours), and average file space in use by project (Pages, 1 page = 512 computer words). These data were accumulated for each project for the months between May, 1985 and May, 1986.

E. H. Shortliffe
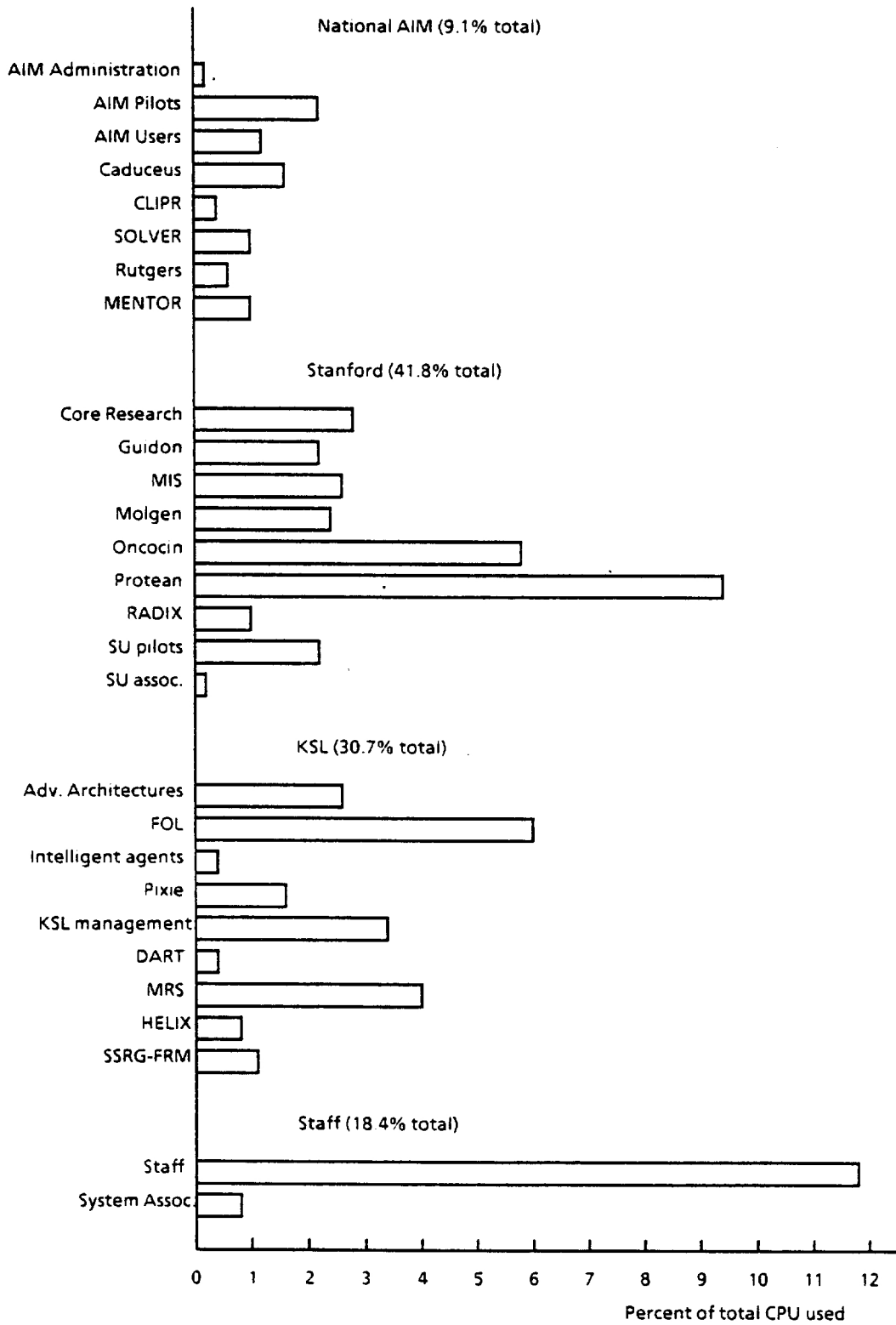
Figure 10:   Cumulative CPU Usage Histogram by Project and Community

Resource Use by Individual Project - 5/85 through 4/86

National AIM Community	CPU (Hours)	Connect (Hours)	File Space (Pages)
1) CADUCEUS	39.32	1274.18	1326

    "Clinical Decision Systems
    Research Resource"
    Jack D. Myers, M.D.
    Harry E. Pople, Jr., Ph.D.
    Randolph A. Miller, M.D.
    University of Pittsburgh
    NIH 5 R24 RR-01101-08
     7/80-6/86 $1,658,347
     7/84-9/85 $354,211
     10/85-3/86 $50,690
    NIH 5 R01 LM03710-05
     7/80-3/86 $853,200
     7/84-9/85 $210,091
     10/85-3/86 $35,316
    NIH New Invest 5 R23 LM03889-03
     Gordon E. Banks, M.D.
     4/82-3/85 $107,675
     4/84-3/85 $35,975
    NIH 1 KO4 LM00084-01
     Randolph A. Miller, M.D.
     9/85-9/86 $55,296
     10/86-9/90   to be determined annually

	CPU (Hours)	Connect (Hours)	File Space (Pages)
2) CLIPR Project	0.75	122.33	144

    "Hierarchical Models
    of Human Cognition"
    Walter Kintsch, Ph.D.
    Peter G. Polson, Ph.D.
    University of Colorado
    NIMH 5 R01 MH-15872-14-16 (Kintsch)
     7/84-6/87 $145,500
     7/83-6/84 $56,501
    NSF (Kintsch)
     8/83-7/86 $200,000(*)
    IBM (Polson)
     David Kieras
     University of Arizona
     1/85-12/86 $500,000
     1/86-12/86 $250,000

3) SOLVER Project	2.06	301.35	650

"Problem Solving
 Expertise"
Paul E. Johnson, Ph.D.
William B. Thompson, Ph.D.
University of Minnesota
Control Data Corp. (Johnson)
 1983-85 $90,000
 1986-88 $95,000
Microelect. and Info. Ctr.
Univ. of MN (Johnson, Thompson,
Slagle, Wechsler, Yonas)
 1984-1985 $500,000
 1985-1986 $300,000
NIH LM-00160 (Johnson, Connelly)
 1984-1989 $712,573
McKnight Foundation (Johnson, Bailey)
 1984-1985 $13,000
Dwan Family Fund, Univ. of MN
Medical School (Johnson)
 1985 $6,000

4) MENTOR Project	27.28	1816.13	650

"Medical Evaluation of Therapeutic
Orders"
Stuart M. Speedie, Ph.D.
University of Maryland
Terrence F. Blaschke, M.D.
Stanford University
National Center for Health
Services Research
1 R18 HS 05263
1/85-12/88 $485,134
1/86-12/86 $182,820

5) *** [Rutgers-AIM] ***			
Rutgers Research Resource	0.03	3.79	0

Artificial Intelligence in Medicine
Casimir Kulikowski, Ph.D.
Sholom Weiss, Ph.D.
Rutgers U., New Brunswick
NIH RR-02230-02 (Kulikowski, Weiss)
 12/83-11/87 $3,198,075
 12/84-11/85 $613,897

6) AIM Pilot Projects	51.92	4754.44	394
7) AIM Administration	0.36	43.92	110
8) AIM Users	27.80	4007.09	2423
	---------	----------	-------
Community Totals	149.52	11838.18	5555

Stanford Community	CPU (Hours)	Connect (Hours)	File Space (Pages)
1) GUIDON-NEOMYCIN Project Bruce G. Buchanan, Ph.D. William J. Clancey, Ph.D. Dept. Computer Science ONR/ARI N00014-79-C-0305 3/79-3/86 $510,311(*) Josiah Macy, Jr. Foundation 3/85-3/88 $503,415	59.02	10173.89	1432
2) MOLGEN Project "Applications of Artificial Intelligence to Molecular Biology: Research in Theory Formation, Testing and Modification" Edward A. Feigenbaum, Ph.D. Peter Friedland, Ph.D. Charles Yanofsky, Ph.D. Depts. Computer Science/ Biology NSF MCS-8310236 (Feigenbaum, Yanofsky) 11/83-10/85 $270,836(*) 11/85-10/86 $135,000(*)	57.66	7121.50	1432
3) ONCOCIN Project "Knowledge Engineering for Med. Consultation" Edward H. Shortliffe, M.D., Ph.D. Dept. Medicine NIH RR-01613 7/83-6/86 $624,455 7/85-6/86 $220,302 NIH LM-04136 8/83-7/86 $211,851 8/85-7/86 $74,150 H.J. Kaiser Family Fdn. 7/83-6/88 $200,000 7/85-6/86 $50,000 NSF IST83-12148 Bruce G. Buchanan (Shortliffe) 3/84-2/87 $330,000(*) 3/85-2/86 $52,679(*) NIH 1 T32 LM07033 7/84-6/89 $903,718 7/85-6/86 $215,850 NIH 1 R23 LM04316 2/85-1/88 $107,441 2/85-1/86 $37,500 NIH 1 R01 LM0442001 9/85-8/88 $314,707 9/85-8/86 $95,205	143.47	242629.39	2619

E. H. Shortliffe

IST 83-12148
 3/84-2/87 $330,000(*)
 3/85-2/86 $52,679

4)	PROTEAN PROJECT Oleg Jardetzky School of Medicine Bruce Buchanan Computer Science Department NSF PCM-84-02348 11/84-10/86 $100,000(*) 11/85-10/86 $50,000(*)	223.52	8577.51	4368
5)	RADIX Project "Deriving Medical Knowledge from Time Oriented Clinical Databases" Robert L. Blum, M.D. Gio C.M. Wiederhold, Ph.D. Depts. Computer Science/ Medicine NSF IST-83-17858 (Blum) 3/84-3/86 $89,597(*) NIH LM-04334 (Wiederhold) 5/84-11/86 $291,192	22.47	2918.85	343
6)	Stanford Pilot Projects	54.25	3468.37	1964
7)	Core AI Research	65.68	8610.18	1263
8)	Stanford Associates	4.19	440.39	1100
9)	Medical Information Sciences	62.19	8227.63	736
		----------	----------	-------
	Community Totals	689.19	74167.62	16777

KSL-AI Community	CPU (Hours)	Connect (Hours)	File Space (Pages)
For funding details please see page 49			
1) Advanced Architectures	64.78	24732.17	1663
2) FOL	142.06	5528.15	1439
3) Intelligent Agent	119.41	13320.77	2840
4) Pixie	35.31	4021.73	875
5) KSL Management	80.02	15484.89	2837
6) DART	11.19	2906.89	811
7) MRS	96.96	8503.16	1456
8) Helix	19.18	7287.03805	
Community totals	504.22	77586.54	14210

SUMEX Staff	CPU (Hours)	Connect (Hours)	File Space (Pages)
1) Staff	280.73	30889.71	9810
2) System Associates	17.83	1235.53	1529
Community Totals	303.27	32686.88	12129

System Operations	CPU (Hours)	Connect (Hours)	File Space (Pages)
1) Operations	752.65	83590.22	9902
Resource Totals	2397.04	279869.50	58573

(*) Award includes indirect costs.

E. H. Shortliffe

System Reliability

System reliability for the DECsystem 2060 has significantly improved in this past period. We have had very few periods of particular hardware or software problems. The data below covers the period of May 1, 1985 to April 30, 1986. The actual downtime was rounded to the nearest hour.

Table 1 : System Downtime Hours per Month - May 1985 through April 1986

8	1	11	9	26	12	1	22	1	7	18	22
May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr

Table 2 : System Downtime Hours per Month - May 1985 through April 1986

Reporting period	:	365 days, 0 hours, 12 minutes, and 49 seconds
Total Up Time	:	358 days, 6 hours, 37 minutes, and 16 seconds
PM Downtime	:	1 days, 8 hours, 49 minutes, and 50 seconds
Actual Downtime	:	5 days, 8 hours, 45 minutes, and 43 seconds
Total Downtime	:	6 days, 17 hours, 35 minutes, and 33 seconds
Mtbf	:	2 days, 14 hours, 45 minutes, and 49 seconds
Uptime Percentage	:	98.53

Network Usage Statistics

The plots in Figure 11 and Figure 12 show the monthly network terminal connect time for the TYMNET and the INTERNET usage. The INTERNET is a broader term for what was previously referred to as Arpanet usage. Since many vendors now support the INTERNET protocols (IP/TCP) in addition to the Arpanet, which converted to IP/TCP in January of 1983, it is no longer possible to distinguish between Arpanet usage and Internet usage on our 2060 system.
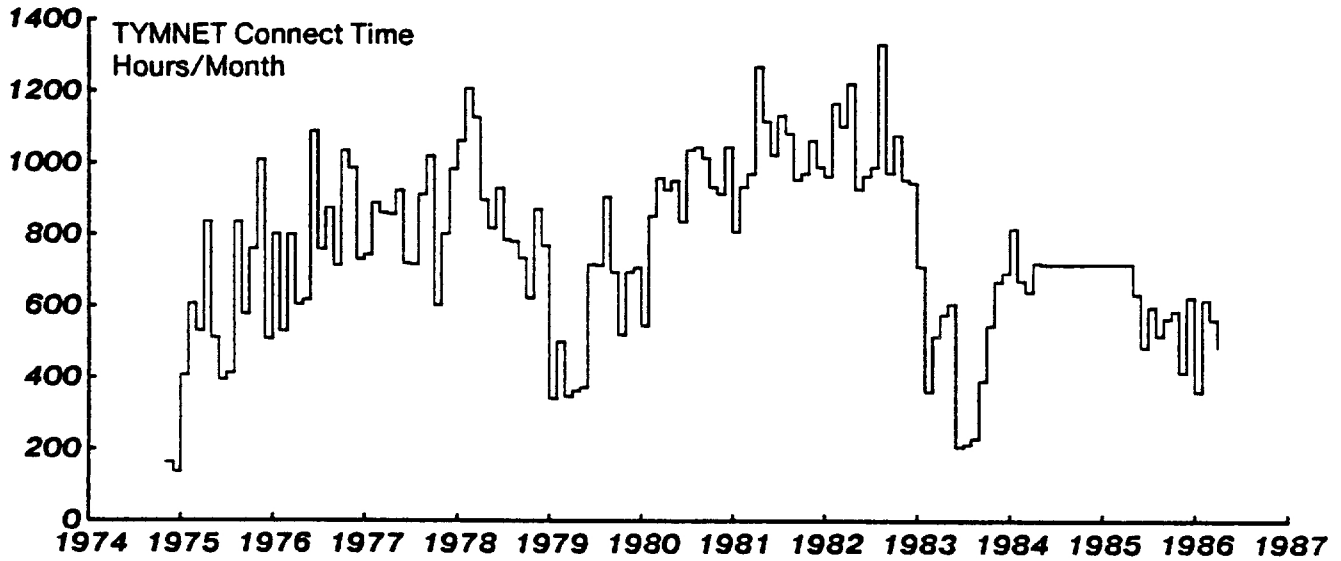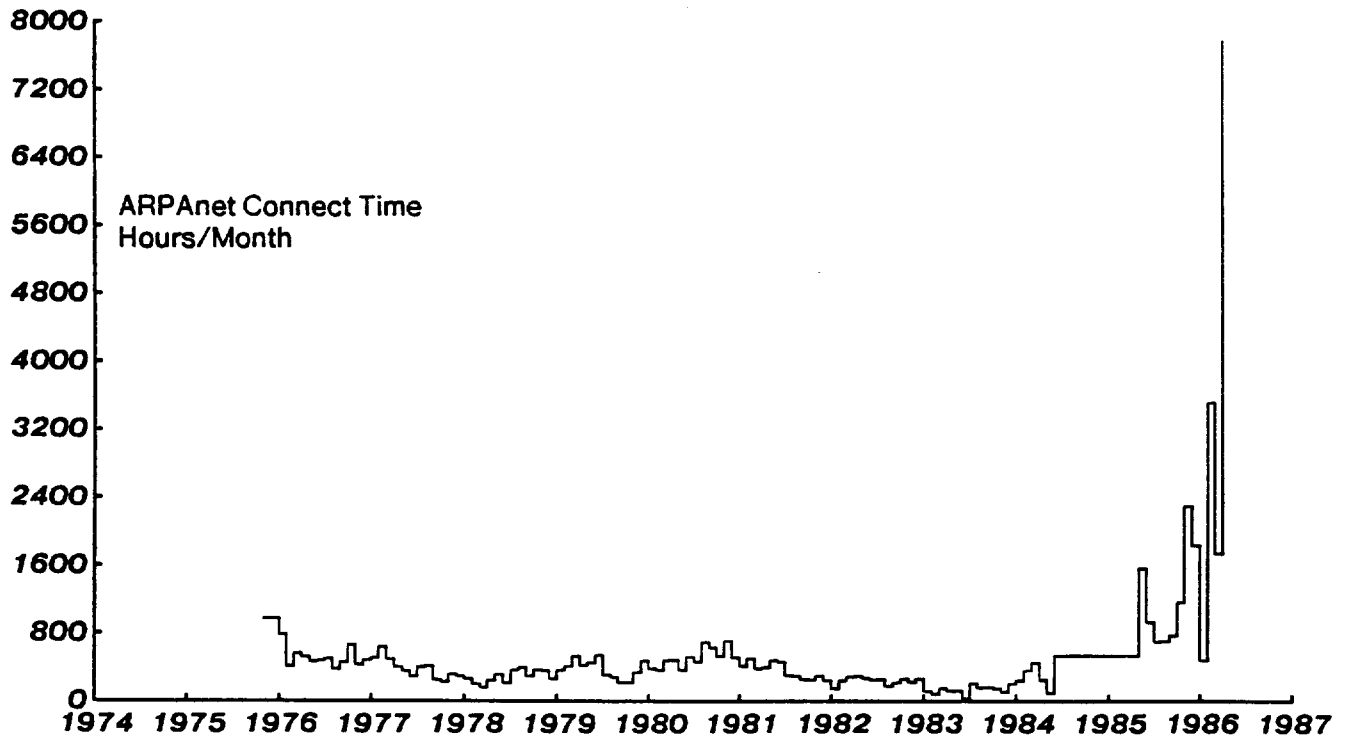
Figure 11: TYMNET Terminal Connect Time

Figure 12: ARPANET Terminal Connect Time

69   E. H. Shortliffe

III.A.4. Future Plans

Our plans for the next grant year (year 14) are based on the Council-approved plans for our 3-year renewal to begin in August, 1986. In the next period, we will merge into SUMEX-AIM two on-going BRTP-supported efforts: the on-going SUMEX-AIM work on biomedical AI applications and the ONCOCIN work on studies in the dissemination of consultation systems. The directions and background for much of this work derive from our research to date.

The goals of the SUMEX-AIM resource are long term in supporting basic research in artificial intelligence, applying these techniques to a broad range of biomedical problems, developing the methodologies for disseminating AI systems into the biomedical community, experimenting with communication technologies to promote scientific interchange, and developing better tools and facilities to carry on this research.

Toward a More Distributed Resource

In the early 1970's, the initial model for SUMEX-AIM as a centralized resource was based on the high cost of powerful computing facilities and the infeasibility of being able to duplicate them readily. As planned, this central role has already evolved significantly and continues to evolve with the introduction of more compact and inexpensive computing technology now available at many more research sites. At the same time, the number of active groups working on biomedical AI problems has grown and the established ones have increased in size. This has led to a growth in the demand for computing resources far beyond what SUMEX-AIM could reasonably and effectively provide on a national scale. We have actively supported efforts by the more mature AIM projects to develop or adapt additional computing facilities tailored to their particular needs and designed to free the main SUMEX resource for new, developing applications projects. To date, over 10 of the national projects have moved some or all of their work to local sites and several have begun resource communities of their own (see page 87). Thus, as more remotely available resources have become established, the balance of the use of the SUMEX-AIM resource has shifted toward supporting start-up pilot projects and the growing AI research community at Stanford.

Summary of Specific Objectives

Our future goals for the central SUMEX-AIM resource are then guided by:

- The increasingly decentralized character of the resource and community and the need to find ways to maintain the scientific communication and sharing that has characterized SUMEX-AIM work

- The continuing exploration of important new areas of biomedical research in which AI techniques can be effectively applied

- The need for a strong basic research effort to investigate methodologies to attack the many problems still beyond our current AI systems and to develop improved tools to build more complex and effective expert systems

- The growing impact of biomedical AI and the need to find and evaluate ways for effectively disseminating biomedical AI technology into real-world settings.

- The need for computing environments for our research and dissemination work that anticipate the needs of AI applications systems over the next 5-10

years, based on the rapidly changing computing hardware and software technology base

SUMEX-AIM will retain its role as a national experimental laboratory for biomedical AI research with a double thrust -- on the one hand, pursuing the basic research for, experimentation on, and trial dissemination of interesting applications and on the other, anticipating and developing the model computing and community environment in which this work can take place. We will nurture existing and new projects and serve as a communications cross-roads for the now diverse AIM community. We will provide the computing resources and some manpower support for long-term basic AI research activities that promise to illuminate issues relevant to future selected collaborative application areas in biology and medicine. For example, our detailed plans have many threads between our basic research in general patient treatment protocol acquisition, representation, and decision-making tools and our collaborative applications in cancer chemotherapy or hypertension trials. Or between our basic research in blackboard problem-solving frameworks and system architectures and our collaborative application in NMR protein conformation determination. Other basic research areas have even longer term goals for problems we hope to be able to address in the future. Underlying all this work will be the development of the Lisp workstation system and network environment that will facilitate these research results and that we feel will become the routine computing environment of the next decade.

In all of this, SUMEX will be both a working laboratory for selected projects within our computing and manpower capacity limits and a source and repository for software and ideas for a broader remote community. We will become an increasingly distributed community resource with heterogeneous computing facilities tethered to each other through various communications media. Many of these machines will be located physically near the projects or biomedical scientists using them. We retain our sincere commitment to our national community of projects. But, inevitably their needs will be met more and more by local facilities and our plans as a resource for the next term place greater emphasis than in the past on supporting the growing Stanford community of AIM collaborations and projects and on developing and integrating model systems at Stanford that can be emulated elsewhere for AIM community needs.

Even with more distributed computing resources, the central resource will continue to play an important role for the next term as a communication crossroads and as a focus for our active dissemination efforts. A key challenge will be to maintain the scientific community ties that grew naturally out of the previous co-location within a central facility.

The following outlines the specific objectives of the SUMEX-AIM resource during the follow-on five year period. Note that these objectives cover only the resource nucleus; objectives for individual collaborating projects are discussed in their respective reports in Section IV. Specific aims are broken into five categories: 1) Core Research and Development, 2) Collaborative Research, 3) Service and Resource Operations, 4) Training and Education, and 5) Dissemination.

1) Core Research and Development

SUMEX funding and computational support for core research is complementary to similar funding from other agencies (see page 49) and contributes to the long-standing interdisciplinary effort at Stanford in basic AI research and expert system design. We expect this work to provide the underpinnings for increasingly effective consultative programs in medicine and for more practical adaptations of this work within emerging microelectronic technologies. Specific aims include:

• Basic research on AI techniques applicable to biomedical problems. Over

the next term we will emphasize work on blackboard problem-solving frameworks and architectures, knowledge acquisition or learning, constraint satisfaction, and qualitative simulation.

- Investigate methodologies for disseminating application systems such as clinical decision-making advisors into user groups. This will include generalized systems for acquiring, representing and reasoning about complex treatment protocols such as are used in cancer chemotherapy and which might be used for clinical trials.

- Support community efforts to organize and generalize AI tools and architectures that have been developed in the context of individual application projects. This will include retrospective evaluations of systems like the AGE blackboard experiment and work on new systems such as BB1, MRS, SOAR, EONCOCIN, EOPAL, Meta-ONYX, and architectures for concurrent symbolic computing. The objective is to evolve a body of software tools that can be used to more efficaciously build future knowledge-based systems and explore other biomedical AI applications.

- Develop more effective workstation systems to serve as the basis for research, biomedical application development, and dissemination. We seek to coordinate basic research, application work, and system development so that the AI software we develop for the next 5-10 years will be appropriate to the hardware and system software environments we expect to be practical by then. Our purchases of new hardware will be limited to experimentation with state-of-the-art workstations as they become available for our system developments.

2) Collaborative Research

- Encourage the exploration of new applications of AI to biomedical research and improve mechanisms for inter- and intra-group collaborations and communications. While AI is our defining theme, we may consider exceptional applications justified by some other unique feature of SUMEX-AIM essential for important biomedical research. We will continue to exploit community expertise and sharing in software development.

- Minimize administrative barriers to the community-oriented goals of SUMEX-AIM and direct our resources toward purely scientific goals. We will retain the current user funding arrangements for projects working on SUMEX facilities. User projects will fund their own manpower and local needs; actively contribute their special expertise to the SUMEX-AIM community; and receive an allocation of computing resources under the control of the AIM management committees. There will be no "fee for service" charges for community members.

- Provide effective and geographically accessible communication facilities to the SUMEX-AIM community for remote collaborations, communications among distributed computing nodes, and experimental testing of AI programs. We will retain the current ARPANET and TYMNET connections for at least the near term and will actively explore other advantageous connections to new communications networks and to dedicated links.

3) Service and Resource Operations

Our applications, core research, and system development will be directed toward realizing and exploiting the computing environment that will be routinely available in the late 1980's and early 1990's, based on compact, decentralized, high-performance personal workstations that take advantage of the intelligent computing environments beginning to emerge from today's Lisp workstations. Consistent with these plans, we will immediately discontinue BRTP subsidy for the DEC 2020 demonstration machine and for the shared VAX 11/780 time-sharing system. Also we will gradually and responsibly phase out BRTP support for the DEC 2060 mainframe system that has been our chief shared resource and link to the past.

SUMEX-AIM does not have the computing or manpower capacity to provide routine service to the large community of mature projects that has developed over the years. Rather, their computing needs are better met by the appropriate development of their own computing resources when justified. Thus, SUMEX-AIM has the primary focus of assisting new start-up or pilot projects in biomedical AI applications in addition to its core research in the setting of a sizable number of collaborative projects. We do offer continuing support for projects through the lengthy process of obtaining funding to establish their own computing base.

Training and Education

- Provide documentation and assistance to interface users to resource facilities and systems.

- Exploit particular areas of expertise within the community for assisting in the development of pilot efforts in new application areas.

- Accept visitors in Stanford research groups within limits of manpower, space, and computing resources.

- Support the Medical Information Science and MS/AI student programs at Stanford to increase the number of research personnel available to work on biomedical AI applications.

- Support workshop activities including collaboration with other community groups on the AIM community workshop and with individual projects for more specialized workshops covering specific research, application, or system dissemination topics.

5) Dissemination

While collaborating projects are responsible for the development and dissemination of their own AI systems and results, the SUMEX resource will work to provide community-wide support for dissemination efforts in areas such as:

- Encourage and support the on-going export of software systems and tools within the AIM community and for commercial development.

- Assist in the production of video tapes and films depicting aspects of AIM community research.

- Promote the publication of books, review papers, and basic research articles on all aspects of SUMEX-AIM research.

Resource Scope

The SUMEX-AIM resource has been from its inception a national experimental resource for biomedical AI with a scope that is carefully defined. Within its limited manpower and computational resources, its focus has been on experiments in new and varied biomedical applications of AI, assisting new research groups in biomedical AI get started, exploring ways to disseminate AI systems into biomedical user communities, supporting relevant basic AI research, and facilitating scientific communications and community sharing. The SUMEX-AIM user community comprises projects from many biological and medical disciplines, ranging from chemistry to molecular biology to clinical medicine to cognitive psychology, and represents collaborations between computer and biomedical scientists from many parts of Stanford University and other universities around the country. The development of this diverse community of projects has both justified the cost of and made effective use of SUMEX-AIM computational and communication facilities at Stanford and elsewhere in our resource community. In its resource role, SUMEX has intentionally limited its production computational capacity to meet the needs of national start-up projects and Stanford research groups, while encouraging self-sufficient community members to develop resources to meet their own computing needs. This has allowed us to provide a level of support for on-going projects and to concentrate most of our efforts on experiments with integrating emerging hardware and software technologies that will be the vehicles of future biomedical AI systems. The results of these experiments are widely disseminated and help other groups through example and direct export of software and ideas.

III.B. Highlights

In this section we describe several research highlights from the past year's activities. These include notes on existing projects that have passed important milestones, new pilot projects that have shown progress in their initial stages, and some other special activities that reflect the impact and influence the SUMEX-AIM resource has had in the scientific and educational communities.

III.B.1. The ONCOCIN Project

The central goal of the ONCOCIN project, under Professor Ted Shortliffe at Stanford Medical School, is to develop programs that can provide advice in oncology protocol management, similar in quality to that given by human experts. The design goals are also to ensure that the system is easy to use and acceptable to physicians. The underlying artificial intelligence research of this project seeks to improve the interactive process, both for the developer of a knowledge-based system, and for the intended end user. In addition, we have emphasized clinical implementation of the developing tool so that we can ascertain the effectiveness of the program's interactive capabilities when it is used by physicians who are caring for patients and are not involved in the computer-based research activity.

ONCOCIN is a large interdisciplinary effort that has involved over 35 individuals since the project's inception in July 1979. The work is currently in its seventh year and significant progress has been made this past year. The OPAL knowledge acquisition system became operational and several new oncology protocols have been entered using this system. As anticipated, we have increased the speed and ease with which protocols can be added to the ONCOCIN knowledge base. Using the protocols entered through OPAL, we have begun experimental testing of the workstation version of ONCOCIN in the Stanford oncology clinic.

We have connected the various parts of the system, and have demonstrated that we have the capability to run ONCOCIN with the reasoning program and interface program on different machines in a communications network. The current version of the program is running on a single workstation, but future versions may take advantage of the multiple machine option. To increase the speed at which we are able to test protocols entered into ONCOCIN, we have developed additional programs to test real and synthetic cases without user interaction; these are then reviewed by our collaborating clinicians.

We have also developed a workstation-based program, OPUS, to help clinicians determine which protocols are appropriate for specific patients. We have been using it in the clinic setting since the end of 1985. Thus, in addition to providing an information resource about protocols, the use of a graphically-oriented program provided a way to learn about the software style and hardware used in the workstation version of ONCOCIN.

The new workstation version of ONCOCIN is still under evaluation. The very promising results of evaluating the older mainframe version was documented in two evaluation papers that appeared in clinical journals[1]. As a further step in our dissemination plans, we are planning experimental installation of ONCOCIN workstations in private oncology offices in San Jose and San Francisco. An application proposing this project is currently under review.

We have also continued our basic research on the design of advanced therapy-planning programs, the ONYX project. We have developed a model for planning which includes techniques from the fields of artificial intelligence, simulation, and decision analysis. Our goal is to allow ONCOCIN to give advice in a wider range of situations; in particular, the system should be able to recommend plans for patients who have an unusual response to chemotherapy.

---

[1] Hickam, D.H., Shortliffe, E.H., Bischoff, M.B., Scott, A.C., Jacobs, C.D. A study of the treatment advice of a computer-based cancer chemotherapy protocol advisor (Memo KSL-85-21). Annals of Internal Medicine 103(6 pt 1):928-936 (1985).

A great deal of interest in ONCOCIN has been shown by the medical, computer science, and lay communities. We are frequently asked to demonstrate the program to Stanford visitors (both the prototype system running in the clinic and the newer work transferring the system to professional workstations). We also demonstrated our workstation system in the Xerox exhibit in the trade show associated with AAAI-84 in Austin, Texas and IJCAI-85 in Los Angeles. Physicians have generally been enthusiastic about ONCOCIN's potential. The interest of the lay community is reflected in the frequent requests for magazine interviews and television coverage of the work. Articles about MYCIN and ONCOCIN have appeared in such diverse publications as *Time* and *Fortune,* and ONCOCIN has been featured on the "NBC Nightly News," the PBS "Health Notes" series, and "The MacNeil-Lehrer Report." Due to the frequent requests for ONCOCIN demonstrations, we have produced a videotape about the ONCOCIN research which includes demonstrations of our professional workstation research projects and the 2020-based clinic system. The tape has been shown at several national meetings, including the 1984 Workshop on Artificial Intelligence in Medicine, the 1984 meeting of the Society for Medical Decision Making, and the 1985 meeting of the Society for Research and Education in Primary Care Internal Medicine. The tape has also been shown to both national and international researchers in biomedical computing. We have also completed an updated tape of our activities for demonstration purposes.

E. H. Shortliffe

III.B.2. The Internist-I Project

The principal objective of Internist-I project, under Professor Jack Myers and Dr. Randolph Miller at the University of Pittsburgh, is to develop a high-level computer diagnostic program in the broad field of internal medicine as an aid in the solution of complex diagnostic problems. To be effective, the program must be capable of multiple diagnoses (related or independent) in a given patient.

A major achievement of this research undertaking has been the design of a program called INTERNIST-I, along with an extensive medical knowledge base. This program has been used over the past decade to analyze many hundreds of difficult diagnostic problems in the field of internal medicine. These problem cases have included cases published in medical journals (particularly Case Records of the Massachusetts General Hospital, in the New England Journal of Medicine), CPCs, and unusual problems of patients in our Medical Center. In most instances, but by no means all, INTERNIST-I has performed at the level of the skilled internist, but this experience has also highlighted several areas of the program needing further improvement.

The development of the QUICK MEDICAL REFERENCE (QMR) system, under the leadership of Dr. Miller, has allowed us to distribute the INTERNIST-I knowledge base in a modified format to over 20 other academic medical institutions. The entire QMR program runs on individual IBM PC-AT workstations. The knowledge base can thereby be used as an "electronic textbook" in medical education at all levels -- by medical students, residents and fellows, and faculty and staff physicians. This distribution is continuing to expand.

The INTERNIST-I program has been used in recent years to develop patient management problems for the American College of Physician's Medical Knowledge Self-assessment Program, and to develop patient management problems and test cases for the Part III Examination and the developing computerized testing program of the National Board of Medical Examiners.

The project staff are continuing to expand the knowledge base of INTERNIST-I and to incorporate the diagnostic consultative program into QMR. The medical knowledge base has continued to grow both in the incorporation of new diseases and the modification of diseases already profiled so as to include recent advances in medical knowledge. Several dozen new diseases have been profiled during the past year. The computer program for the interrogative part of the diagnostic program is the main remaining task. An editor for the QMR knowledge base, as modified from the INTERNIST-I knowledge base, is nearing completion.

In the near future the project will be ready to incorporate into the QMR diagnostic consultant program the enhancements of the INTERNIST-I knowledge base, e.g. the use of "facets" of diseases or syndromes. This addition is expected to improve the performance of the diagnostic consultant program.

III.B.3. The PROTEAN Project

The goals of this project, under Professors Oleg Jardetzky and Bruce Buchanan at Stanford University, are related both to biochemistry and artificial intelligence: (a) use existing AI methods to aid in the determination of the 3-dimensional structure of proteins in solution (as opposed to crystallized proteins), and (b) use protein structure determination as a test problem for experiments with the AI problem solving structure known as the Blackboard Model. Empirical data from nuclear magnetic resonance (NMR) and other sources may provide enough constraints on structural descriptions to allow protein chemists to bypass the laborious methods of crystallizing a protein and using X-ray crystallography to determine its structure. This problem exhibits considerable complexity. Yet there is reason to believe that AI programs can be written that reason much as experts do to resolve these difficulties.

The molecular structure of proteins is essential for understanding many problems of medicine at the molecular level, such as the mechanisms of drug action. Using NMR data from proteins in solution will allow the study of proteins whose structure cannot be determined with other techniques, and will decrease the time needed for the determination.

Over the past year, we have constructed a prototype of such a program, called PROTEAN, designed on the blackboard model. It is implemented in BB1, a framework system for building blackboard systems that control their own problem-solving behavior. The reasoning component of PROTEAN directs the actions of the Geometry System (GS), a set of programs that performs the computationally intensive task of positioning portions of a molecule with respect to each other in three dimensions. Currently we have implemented two versions of the GS: an InterLISP version used for quickly testing ideas and developing prototypes of geometric routines on a LISP workstation; and a high performance version written in C and running in the UNIX environment, providing efficient computations on a VAX 11/780. We have coupled the reasoning and geometry programs with an IRIS graphics terminal (shared with SUMEX) which displays the evolving protein structures at several levels of detail. The display in three dimensions is essential to understanding the behavior of the reasoning and geometry systems, and provides essential insights on the problem solving process.

The current version of PROTEAN has five domain knowledge sources and five control knowledge sources that demonstrate the reasoning techniques described above for the assembly of a protein. Each domain knowledge source directs a small portion of the construction of the molecule. These knowledge sources develop partial solutions that position alpha helices, beta strands, and random coils at the Solid level and refine the resulting state families using all available distance constraints. The control knowledge sources determine which of the possible assembly actions is the best to perform at each stage of the problem solving.

The PROTEAN system [2] has been used to construct a complete solution at the solid level of detail for the Lac-repressor headpiece, a protein with 51 amino acids consisting of 4 random coil sections and three alpha helices[1]. In this work, the constraints were determined experimentally from NMR studies.

To demonstrate that our method is correct, we have applied PROTEAN to sperm whale myoglobin, a molecule whose crystal structure is known. By using distances between atoms in the crystal, distance constraints were applied to the eight helices in myoglobin

[1]Buchanan, B.G., Altman, A., Brinkley, J., Cornelius, C., Duncan, B., Hayes-Roth, B., Hewett, M., Lichtarge, O., Jardetzky, O.: *The Heuristic Refinement Method for Deriving Solution Structures of Proteins.* Report KSL-85-41. October 1985. In *Proceedings of the National Academy of Sciences.*.

to determine if PROTEAN would reproduce the crystal structure. The family of solutions obtained from PROTEAN includes the actual structure of the molecule. Work is proceeding to include the heme group of myoglobin as a component and use constraints to other portions of the molecule to further restrict the state families obtained.

III.B.4. AIM Community Software Support

The SUMEX-AIM resource staff have made major efforts to assist other members of the SUMEX-AIM community in integrating the technologies needed for biomedical AI research. This is often achieved through direct contact with staff members at these institutions, at meetings and workshops, or via electronic mailing lists. For example, the SUMEX-AIM gateway software is well known for its versatility and reliability, and has been distributed to both the University of Chicago and Rutgers University through contacts made this way. The gateway running at Rutgers is quite similar to our own and uses Ethernet as its medium of transport. The gateway at the University of Chicago is a principle node on a fibre-optic network.

Dr. Charles Hedrick, Director of the Rutgers Computer Science Department computer facility, acknowledged our contribution to the Rutgers networking efforts as follows:

SUMEX Support of Networking at Rutgers

During the last two years, computing at Rutgers has become increasingly network-oriented. This change has included the Rutgers AIM project, as well as most other compute-intensive research areas within the University. Our primary computing resource now consists of workstations such as Suns or MicroVAXes for individual researchers, and super-minis at the departmental or project level. All of these systems are connected to each other via Ethernet. The workstations depend upon the super-minis for a number of services, including file storage, backup, and printing. In addition, there are more specialized services available via the Ethernet, including larger mainframes, a high-speed Xerox laser printer, and (shortly) the facilities of the John von Neumann Supercomputer Center. The Ethernet-based network has become critical to the operation of most of our computing facilities.

To a large extent the network technology now used at Rutgers is based on work done at Stanford, and particularly at SUMEX. The staff at SUMEX have given us significant assistance in importing this technology to Rutgers. Much of the AI research at Rutgers is done using Xerox Interlisp-D workstations. These workstations require network file service via either the PUP or XNS protocols. Our primary file server is a Pyramid Unix system that uses the Unix PUP implementation developed at SUMEX.

During the last year, the amount of Ethernet work at Rutgers became large enough that a single Ethernet was no longer sufficient. At that point we created 4 separate Ethernets. They are connected with an Ethernet gateway developed at SUMEX. SUMEX personnel supplied us with a description of the hardware necessary. It could be assembled from commercially-available pieces. They also supplied us with the software to be run in the gateway, and the development tools necessary to compile and down-load it. Although our setup is very similar to the one at Stanford, a certain amount of customization was necessary for the environment here. Most of this work was done by the SUMEX staff. They also provided assistance for the rest. We have the highest praise for the quality of this software, and for the support that we have gotten from the SUMEX staff, particularly Bill Yeager. This is the only piece of software we have that has never crashed or otherwise misbehaved. Its overall quality, including such items as administrative tools, is better than most commercially-supported products we have seen.

We consider this sort of research and technology transfer to be an important benefit of SUMEX. The tendency now is for each research group to have its own equipment and its own support staff. This is both inevitable and useful. However it is important for the AIM community to have an organization such as

SUMEX that can investigate new computing technology and make it available to the rest of the community.

III.B.5. Remote Virtual Graphics

Lisp workstations of various types have proven extremely powerful, both as development environments for artificial intelligence research and as vehicles for disseminating AI systems into user communities. In addition to the compact, inexpensive computing resources workstations provide, high-quality graphics play a key role in their power. Such graphics systems have become indispensable for understanding the complex data structures involved in developing and debugging large AI systems and are important in facilitating user access to working programs (e.g., for ONCOCIN and PROTEAN). However, as we move towards a distributed workstation computing environment for AI research in the SUMEX-AIM community (and move away from the centralized, shared DEC 2060), a number of technical obstacles must be overcome. One of the most important is to eliminate the need for the workstation display to be situated close to the workstation computing engine.

In the past, members of the SUMEX-AIM community have often watched each others programs work by linking their CRT terminals to the text output of a running program on the SUMEX 2060. In the case of workstations, though, it is much more difficult to link across several networks to view the complex graphics output of a program. Even locally, it is important to make graphical interaction with workstations across campus or from home possible. One would like to be able to provide the same powerful graphical tools and programming environment that are available to a user sitting in front of the workstation to the remote user if that user has a low cost bit-mapped display and mouse. In order to accomplish this, it is necessary to capture and encode the many graphics operations involved so that they can be sent over a relative low-speed network connection with the same interactive facility as if one had the display connected through the dedicated high-speed (30 Mhz) native vendor display/workstation connection.

In order to study the feasibility of remote access to workstations to satisfy these requirements, we have been experimenting with virtual graphics protocols (VGP) to capture graphics constructs and interactions at a high level. Such a protocol, the Virtual Graphics Terminal Service (VGTS), was developed at Stanford in the Computer Science distributed systems group[1]. The VGTS provides tools to define objects like windows, lines, rectangles, circles, bitmaps, ellipses, splines, and graphics events like mouse clicks independently of the graphics hardware and operating systems. This encoding minimizes the communication bandwidth required between cooperating hosts, for example, to remotely draw a line.

Over the past year, an implementation of this protocol was developed and installed in the operating system of a Xerox 1186 Lisp workstation so that its presence would be transparent to the programmer. This means that if one connects to such a LISP workstation from a SUN workstation (running suitable VGTS software), the Lisp machine graphics will be sent over the net and reconstructed on the SUN workstation without changes to the application program running. This implementation has worked very well in early experiments so that over an Ethernet, the remote response time is quite close to the response time on the Lisp machine itself.

As a consequence of this work, we have demonstrated the feasibility of remotely using LISP workstations over an Ethernet to take advantage of their graphics programming environment. Work in the virtual graphics group at Stanford also suggests that the VGTS can be used effectively over the ARPANET (56 KBits/sec), at least for some kinds of applications, with excellent response time. Much more work needs to be done

---

[1] see Keith A. Lantz, David R. Cheriton, and William I. Nowicki, *Third Generation Graphics for Distributed Systems*, Computer Systems Laboratory Report, Stanford University, December 21, 1982.

in this area to fully develop and integrate these capabilities into Lisp machine systems and to insure that cross-country connections will indeed give usable response time. Success of this work will mean that one can use LISP machine systems from TYMNET, UNINET, ARPANET, or an Ether TIP connection throughout the SUMEX-AIM community.

III.C. Administrative Changes

There have been few administrative changes within the project this past year other than some turn-over in staff. In the coming year, Professor Shortliffe will be on sabbatical at the University of Pennsylvania. During this time, Professor Feigenbaum will reassume the role of principal investigator.

More substantial developments are expected in the coming year as well. First we will be moving the Medical Computer Science and SUMEX offices into newly constructed space within a Stanford Medical School Office Building. We will occupy approximately 6500 square feet which almost doubles the space available to us. This space is designed to improve the interactions within our groups.

The other development will be the design of a cost recovery system as we phase out BRTP subsidy of the DEC 2060 facility. We plan to use the continuing component of NIH support to retain our no fee for service policy for non-Stanford projects for as long as possible. We will establish a cost center to allocate the uncovered expenses to Stanford user projects.

III.D. Resource Management and Allocation

Early in the design of the SUMEX-AIM resource, an effective management plan was worked out with the Biotechnology Resources Program (now Biomedical Research Technology Program) at NIH to assure fair administration of the resource for both Stanford and national users and to provide a framework for recruitment and development of a scientifically meritorious community of application projects. This structure has been described in some detail in earlier reports and is documented in our recent renewal application. It has continued to function effectively as summarized below.

- The AIM Executive Committee meets regularly by teleconference to advise on new project applications, discuss resource management policies, plan workshop activities, and conduct other community business. The Advisory Group meets together at the annual AIM workshop to discuss general resource business and individual members are contacted much more frequently to review project applications. (See Appendix A on page 187 for a current listing of AIM committee membership).

- We have actively recruited new application projects and disseminated information about the resource. The number of formal projects in the SUMEX-AIM community still runs at the capacity of our computing resources. With the development of more decentralized computing resources within the AIM community outside of Stanford (see below), the center of mass of our community has naturally shifted toward the growing number of Stanford applications and core research projects. We still, however, actively support new applications in the national community where these are not able to gain access to suitable computing resources on their own.

- With the advice of the Executive Committee, we have awarded pilot project status to promising new application projects and investigators and where appropriate, offered guidance for the more effective formulation of research plans and for the establishment of research collaborations between biomedical and computer science investigators.

- We have carefully reviewed on-going projects with our management committees to maintain a high scientific quality and relevance to our biomedical AI goals and to maximize the resources available for newly developing applications projects. Several fully authorized and pilot projects have been encouraged to develop their own computing resources separate from SUMEX or have been phased off of SUMEX as a result and more productive collaborative ties established for others.

- We have continued to provide active support for the AIM workshops. The last one was held in Washington, DC, hosted by the National Library of Medicine under Drs. Lindberg and Kingsland.

- We will have to partly abandon our policy of no fee-for-service for projects using the SUMEX resource. This policy had effectively eliminated the serious administrative barriers that would have blocked our research goals of broader scientific collaborations and interchange on a national scale within the selected AIM community. We will attempt to minimize these barriers for national projects but will be obliged to recover costs not supported by BRTP as we phase out support for our mainframe computers (DEC 2060, 2020, and VAX 11/780).

- We have tailored resource policies to aid users whenever possible within our research mandate and available facilities. Our approach to system scheduling, overload control, file space management, etc. all attempt to give users the greatest latitude possible to pursue their research goals consistent with fairly meeting our responsibilities in administering SUMEX as a national resource.

As indicated above, we have sought to retain SUMEX resources for new projects, those exploring new areas in biomedical AI applications and those in such an early state of feasibility that they are unable to afford their own computing resources. This policy has worked effectively as seen from the following lists of terminated projects and projects now using their own computing resources at other sites:

Projects Moved All or In Part to Other Machines:

Stanford Projects:

- GENET [Brutlag, Kedes, Friedland – IntelliCorp]

National Projects:

- Acquisition of Cognitive Procedures (ACT) [Anderson – CMU]

- Chemical Synthesis [Wipke – UC Santa Cruz]

- Simulation of Cognitive Processes [Lesgold – Pittsburgh]

- PUFF [Osborne, Feigenbaum, Fagan – Pacific Medical Center]

- CADUCEUS/INTERNIST [Pople, Myers – Pittsburgh]

- Rutgers [Amarel, Kulikowski, Weiss – Rutgers]

- MDX [Chandrasekaran – Ohio State]

- SOLVER [P. Johnson – University of Minnesota]

Completed Projects Summary

Stanford Projects:

- DENDRAL [Lederberg, Djerassi, Buchanan, Feigenbaum]

- MYCIN [Shortliffe, Buchanan]

- EMYCIN [Shortliffe, Buchanan]

- CRYSALIS [Feigenbaum, Engelmore]

- MOLGEN I [Feigenbaum, Brutlag, Kedes, Friedland]

- AI Handbook [Feigenbaum, Barr, Cohen]

- AGE Development [Feigenbaum, Nii]

National Projects:

- Ventilator Management [Osborne, Feigenbaum, Fagan – Pacific Medical Center]

- Higher Mental Functions [Colby - USC]

III.E. Dissemination of Resource Information

Throughout the history of the SUMEX-AIM resource, we have made extensive efforts at disseminating the AI technology developed here. This has taken the form of many publications -- over 45 combined books and papers are published per year from the KSL; wide distribution of our software including systems software and AI application and tool software, both to other research laboratories and for commercial development; production of films and video tapes depicting aspects of our work; and significant project efforts at studying the dissemination of individual applications systems such as the GENET community (DNA sequence analysis software) and the ONCOCIN resource-related research project (see 109).

Software Distribution

We have widely distributed both our system software and our AI tool software. We have no accurate records of the extent of distribution of the system codes because their distribution is not centralized and controlled. The recent programs such as TOPS-20 monitor enhancements, the Ethernet gateway and TIP programs, the SEAGATE AppleBus to Ethernet gateway, the PUP Leaf server, the SUMACC development system for Macintosh workstations, and our Lisp workstation programs are well-distributed throughout the ARPANET community and beyond.

We have made major efforts to assist other members of the SUMEX-AIM community in integrating the technologies needed for biomedical AI research. This is often achieved through direct contact with staff members at these institutions at meetings and workshops or via electronic mailing lists. For example, the SUMEX-AIM gateway software is well known for its versatility and reliability, and has been distributed to both the University of Chicago and Rutgers University through contacts made this way. The gateway running at Rutgers is quite similar to our own and uses Ethernet as its medium of transport. The gateway at the University of Chicago is a principle node on a fibre-optic network. Dr. Charles Hedrick, Director of the Rutgers Computer Science Department facility, acknowledged our contribution to the Rutgers networking efforts as follows:

SUMEX Support of Networking at Rutgers

During the last two years, computing at Rutgers has become increasingly network-oriented. This change has included the Rutgers AIM project, as well as most other compute-intensive research areas within the University. Our primary computing resource now consists of workstations such as Suns or MicroVAXes for individual researchers, and super-minis at the departmental or project level. All of these systems are connected to each other via Ethernet. The workstations depend upon the super-minis for a number of services, including file storage, backup, and printing. In addition, there are more specialized services available via the Ethernet, including larger mainframes, a high-speed Xerox laser printer, and (shortly) the facilities of the John von Neumann Supercomputer Center. The Ethernet-based network has become critical to the operation of most of our computing facilities.

To a large extent the network technology now used at Rutgers is based on work done at Stanford, and particularly at SUMEX. The staff at SUMEX have given us significant assistance in importing this technology to Rutgers. Much of the AI research at Rutgers is done using Xerox Interlisp-D workstations. These workstations require network file service via either the PUP or XNS protocols. Our primary file server is a Pyramid Unix system that uses the Unix PUP implementation developed at SUMEX.

During the last year, the amount of Ethernet work at Rutgers became large enough that a single Ethernet was no longer sufficient. At that point we created 4 separate Ethernets. They are connected with an Ethernet gateway developed at SUMEX. SUMEX personnel supplied us with a description of the hardware necessary. It could be assembled from commercially-available pieces. They also supplied us with the software to be run in the gateway, and the development tools necessary to compile and down-load it. Although our setup is very similar to the one at Stanford, a certain amount of customization was necessary for the environment here. Most of this work was done by the SUMEX staff. They also provided assistance for the rest. We have the highest praise for the quality of this software, and for the support that we have gotten from the SUMEX staff, particularly Bill Yeager. This is the only piece of software we have that has never crashed or otherwise misbehaved. Its overall quality, including such items as administrative tools, is better than most commercially-supported products we have seen.

We consider this sort of research and technology transfer to be an important benefit of SUMEX. The tendency now is for each research group to have its own equipment and its own support staff. This is both inevitable and useful. However it is important for the AIM community to have an organization such as SUMEX that can investigate new computing technology and make it available to the rest of the community.

We do have reasonably accurate records of the distribution of our AI tool software because the recipient community is more directly coupled to us and the distribution is centralized. Records indicate that over the past three years there have been about 1000 inquiries that have resulted in the distribution of written material about our software systems. It is likely that there have been a similar number of unrecorded or informal interactions on the part of the staff.

In the past year, we have distributed 73 copies of various software packages, including 18 EMYCIN systems, 41 MRS systems, 7 AGE systems, and 7 BB1 systems. In addition, a substantial amount of time was spent organizing the 106 files of the three volumes of the A/I Handbook so that it could be provided to two collaborators. And finally, we are organizing the DENDRAL software to make it available for distribution again. The CONGEN portion is essentially complete and being tested by an outside user and one of the original developers. Some progress has been made with the organizing of the GENOA portion. Several interested groups are awaiting its release.

Video Tapes and Films

The KSL and the ONCOCIN project have prepared several video tapes that provide an overview of the research and research methodologies underlying our work and that demonstrate the capabilities of particular systems. These tapes are available through our groups, the Fleischmann Learning Center at the Stanford Medical Center, and the Stanford Computer Forum and copies have been mailed to program offices of our various funding sponsors. The three tapes include:

- *Knowledge Engineering in the Heuristic Programming Project* -- This 20-minute film/tape illustrates key ideas in knowledge-based system design and implementation, using examples from ONCOCIN, PROTEAN, and knowledge-based VLSI design systems. It describes the research environment of the KSL and lays out the methodologies of our work and the long term research goals that guide it.

- *ONCOCIN Overview* -- This is a 30-minute tape providing an overview of the ONCOCIN project. It gives an historical context for the work, discusses

the clinical problem and the setting in which the prototype system is being used, and outlines the plans for transferring the system to run on single-user workstations. Brief illustrations of the graphics capabilities of ONCOCIN on a Lisp workstation are also provided.

- *ONCOCIN Demonstration* -- This 1-hour tape provides detailed examples of the key components of the ONCOCIN system. It begins with a demonstration of the prototype system's performance on a time-shared mainframe computer and then shows each of the elements involved in transferring the system to Lisp workstations.

III.F. Suggestions and Comments

Resource Organization

We continue to believe that the Biomedical Research Technology Program is one of the most effective vehicles for developing and disseminating technological tools for biomedical research. The goals and methods of the program are well-designed to encourage building of the necessary multi-disciplinary groups and merging of the appropriate technological and medical disciplines.

Electronic Communications

SUMEX-AIM has pioneered in developing more effective methods for facilitating scientific communication. Whereas face-to-face contacts continue to play a key role, in the longer-term we feel that computer-based communications will become increasingly important to the NIH and the distributed resources of the biomedical community. We would like to see the BRTP take a more active role in promoting these tools within the NIH and its grantee community.

IV. Description of Scientific Subprojects

The following subsections report on the AIM community of projects and "pilot" efforts including local and national users of the SUMEX-AIM facility at Stanford. However, those projects admitted to the National AIM community which use the Rutgers-AIM resource as their home base are not explicitly reported here.

In addition to these detailed progress reports, abstracts for each project and its individual users are submitted on a separate Scientific Subproject Form. However, we have included here briefer summary abstracts of the fully-authorized projects in Appendix B on page 191.

The collaborative project reports and comments are the result of a solicitation for contributions sent to each of the project Principal Investigators requesting the following information:

I. SUMMARY OF RESEARCH PROGRAM
   A. Project rationale
   B. Medical relevance and collaboration
   C. Highlights of research progress
      --Accomplishments this past year
      --Research in progress
   D. List of relevant publications
   E. Funding support

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE
   A. Medical collaborations and program dissemination via SUMEX
   B. Sharing and interactions with other SUMEX-AIM projects
      (via computing facilities, workshops, personal contacts, etc.)
   C. Critique of resource management
      (community facilitation, computer services, communications
      services, capacity, etc.)

III. RESEARCH PLANS
   A. Project goals and plans
      --Near-term
      --Long-range
   B. Justification and requirements for continued SUMEX use
   C. Needs and plans for other computing resources beyond SUMEX-AIM
   D. Recommendations for future community and resource development

We believe that the reports of the individual projects speak for themselves as rationales for participation. In any case, the reports are recorded as submitted and are the responsibility of the indicated project leaders. The only exceptions are the respective lists of relevant publications which have been uniformly formatted for parallel reporting on the Scientific Subproject Form.

IV.A. Stanford Projects

The following group of projects is formally approved for access to the Stanford aliquot of the SUMEX-AIM resource. Their access is based on review by the Stanford Advisory Group and approval by Professor Feigenbaum as Principal Investigator.

In addition to the progress reports presented here, abstracts for each project and its individual users are submitted on a separate Scientific Subproject Form.

IV.A.1. GUIDON/NEOMYCIN Project

GUIDON/NEOMYCIN Project

William J. Clancey, Ph.D.
Department Computer Science
Stanford University

Bruce G. Buchanan, Ph.D.
Computer Science Department
Stanford University

I.   SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

The GUIDON/NEOMYCIN Project is a research program devoted to the development of a knowledge-based tutoring system for application to medicine.   This work derived from our first system, the MYCIN program.   That research led to three sub-projects (EMYCIN, GUIDON, and ONCOCIN) described in previous annual reports.   EMYCIN has been completed and its resources reallocated to other projects.   GUIDON and ONCOCIN have become projects in their own right.

The key issue for the GUIDON/NEOMYCIN project is to develop a program that can provide advice similar in quality to that given by human experts, modeling how they structure their knowledge as well as their problem-solving procedures.   The consultation program using this knowledge is called NEOMYCIN.   NEOMYCIN's knowledge base, designed for use in a teaching application, is the subject material used by a family of instructional programs referred to collectively as GUIDON2.   The problem-solving procedures are developed by running test cases through NEOMYCIN and comparing them to expert behavior.   Also, we use NEOMYCIN as a test bed for the explanation capabilities incorporated in our instructional programs.

The purpose of the current contracts is to construct a knowledge-based tutoring system that teaches diagnostic strategies explicitly.   By strategy, we mean plans for establishing a set of possible diagnoses, focusing on and confirming individual diagnoses, gathering data, and processing new data.   The tutorial program has capabilities to recognize these plans, as well as to articulate strategies in explanations about how to do diagnosis. The strategies represented in the program, modeling techniques, and explanation techniques are wholly separate from the knowledge base, so that they can be used with many medical (and non-medical) domains.   That is, the target program will be able to be tested with other knowledge bases, using system-building tools that we provide.

B. Medical Relevance and Collaboration

There is a growing realization that medical knowledge, originally codified for the purpose of computer-based consultations, may be utilized in additional ways that are medically relevant.   Using the knowledge to teach medical students is perhaps foremost among these, and NEOMYCIN continues to focus on methods for augmenting clinical knowledge in order to facilitate its use in a tutorial setting.   A particularly important aspect of this work is the insight that has been gained regarding the need to structure knowledge differently, and in more detail, when it is being used for different purposes (e.g., teaching as opposed to clinical decision making).   It was this aspect of the

GUIDON research that led to the development of NEOMYCIN, which is an evolving computational model of medical diagnostic reasoning that we hope will enable us to better understand and teach diagnosis to students. An important additional realization is that these structuring methods are beneficial for improving the problem-solving performance of consultation programs, providing more detailed and abstract explanations to consultation users, and making knowledge bases easier to maintain.

As we move from technological development of explanation and student modeling capabilities, we are now collaborating closely with medical students and physicians to design an effective, useful tutoring program. Stanford Medical School faculty, such as Dr. Maffly, have shown considerable interest in this project. A research fellow associated with Maffly, Curt Kapsner, M.D., worked with the project for two years while at Stanford. The project also collaborates with a community of researchers focusing on medical education, funded by the Josiah Macy, Jr. Foundation.

C. Highlights of Research Progress

C.1 Accomplishments This Past Year

C.1.1 The NEOMYCIN Consultation Program

This year our physician and medical student exhaustively reviewed the knowledge base, filling in gaps and improving entries. Disease terminology and rule translation were modified, so that our medical collaborators now believe that the system is ready for use by students.

C.1.2 The GUIDON-WATCH Display Program

In the past few years, we have developed a complex interface for viewing NEOMYCIN's knowledge base and reasoning process. We call this program GUIDON-WATCH; it serves as the foundation for our teaching programs.

Major improvements were made this year to the display interface and underlying code. Interfacing a knowledge base with a bit-map display requires new programming methods. We standardized techniques for describing display structures (hierarchies, tables, and menus) and relating them to the various display operations (resizing text, highlighting subtrees, getting details, making hardcopy). We are teaching these coding methods to our new students, significantly improving their productivity and the clarity of the resulting program.

The display system for showing NEOMYCIN's reasoning was totally revamped to use multiple fonts for clarity, to allow uniform "buttoning" of any item to get additional information, and to simplify printout for a student. The simplification took several months in late 1985 and has been used by medical students in informal trials.

C.1.3 The GUIDON-DEBUG Instructional Program

During the past year, we developed a prototype version of GUIDON-DEBUG. Briefly, a student plays the role of a knowledge engineer, running cases and editing the knowledge base in order to find the missing or incorrect knowledge we have planted. The patient-specific model focuses his search, revealing what findings are not accounted for, suggesting what rules might be missing.

We provide extensive graphic facilities (GUIDON-WATCH), so the student needs to know little about the underlying implementation. Indeed, the idea is to make it easy to learn how NEOMYCIN works, rather than to hide what we are trying to teach. We have written about the theory of learning and novelty of our approach in publications cited elsewhere.

The idea of the *patient-specific model* window is the most significant technical advance

in the past few months. This window shows the support for each disease hypothesis as a kind of proof. It quickly reveals gaps in the diagnostic explanation provided by the program, in a way that is hidden by the accumulating of evidence or scoring scheme used by most programs (including NEOMYCIN before we made this change). The idea of constructing this window comes from a mixture of sources: Patil's idea of a patient-specific model, Reggia's idea of "coverage" of disease hypotheses, Anderson's idea of showing a geometry proof as a graph, and Brown's idea that we should use graphics to help students visualize the reasoning process. Everyone, particularly our medical collaborators, has rallied around this conceptualization.

We have integrated our student modeling program, ODYSSEUS, into the GUIDON-DEBUG system. The modeling system is given a reordered set of data requests, with perhaps additional requests, which it parses with respect to NEOMYCIN's diagnostic strategy. That is, it attempts to find a strategic explanation for why each question was asked, abstracting coherent sequences from the list of data requests.

A new program relates this parse to the knowledge base, determining what difference to the knowledge base would produce the indicated ordering of questions. This serves both as an evaluation of a student's critique, as well as a very efficient knowledge acquisition device. It is gratifying to have this capability implemented, since it is the original idea that motivated the original design of NEOMYCIN.

An interesting capability was implemented that allows the user (student, physician, or knowledge engineer) to show the state of the consultation program at any time. In GUIDON-DEBUG you underline a question asked by NEOMYCIN and then display any window you wish. For example, if you say WHY, the program will answer the question as if you had asked it at that time during the consultation.

C.1.4 Dissemination of results

In the past year, we had a number of outstanding opportunities to present our work to other researchers. Dr. Clancey was an invited speaker at several European conferences, giving forty hours of lectures in Belgium, England, France, Germany, Holland, Italy, and Switzerland during this period. The above programs were demonstrated in five of these countries, as well as receiving strong attention at IJCAI in August in Los Angeles. (Demonstrations were given on Xerox Dandelions supplied by local representatives of Rank-Xerox. We are impressed by how easy it is to merely carry floppies to Europe and simply start up the program on the Xerox personal workstation.)

As the program becomes ready for demonstrating in tutorial mode, we are showing it to more people with a medical background. In January, Dr. Clancey gave a lecture and demonstration at a seminar run by the AAMC in San Diego (Information Management Technology). The program was also demonstrated at a Biomedical conference. We also entertained G. Octo Barnett for a morning, introducing him to the idea of intelligent tutoring systems. We are continuing to work with one of Dr. Maffly's medical students; Maffly is the main proponent of CAI in the Stanford Medical School. This student carries our work back to Maffly's lab, beginning the process of integrating our research with more traditional CAI practice.

C.2 Research in Progress

The following projects are active as of June 1986 (see also near-term plans listed in Section III.A):

1. Developing additional instructional programs based on NEOMYCIN;

2. Studying learning in the setting of debugging a knowledge base;

3. Re-implementing the explanation program to use the logic-encoding of the

metarules (stating this program in the same task/metarule language so that it might reason about its own explanations);

4. Generalizing our graphics package using object-oriented techniques;

5. Applying the student modeling program, ODYSSEUS, to knowledge acquisition; and

6. Preparing HERACLES, the generalization of NEOMYCIN, for use by other people.

D. Publications Since January 1985

1. (*) Clancey, W.J.: *Review of Sowa's "Conceptual Structures: Information Processing in Mind and Machine." Journal of Artificial Intelligence* 27(1):113-124, 1985.

2. (*) Clancey, W. J.: *Expert system design: Lessons from applications to education (abstract)*. IEEE Spring Computer Conference, San Francisco, page 115.

3. (*) Thompson, T. and Clancey, W. J.: *A qualitative modeling shell for process diagnosis. IEEE Software* 3(2):6-15, March 1986.

4. (*) Clancey, W.J.:*The science and engineering of qualitative models.* Submitted to AAAI-86, March 1986.

5. (*) Clancey, W.J., Richer, M., Wilkins, D.C., Barnhouse, S., Kapsner, C., Leserman, D., Macias, J., Merchant, A., and Rodolitz, N.: *Guidon-Debug: The student as knowledge engineer.* Submitted to AAAI-86, April 1986.

6. (*) Clancey W.J.: *Qualitative Student Models.* **Annual Review of Computer Science.** Palo Alto: Annual Reviews, Report KSL-86-11, Computer Science Dept., May 1986.

7. (*) Clancey W.J.: *From GUIDON to NEOMYCIN and HERACLES in twenty short lessons: ONR Final Report, 1979-1985.* To appear in the *AI Magazine,* August 1986.

E. Funding Support

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaborations and Program Dissemination via SUMEX

A great deal of interest in GUIDON and NEOMYCIN has been shown by the medical and computer science communities. We are frequently asked to demonstrate these programs to Stanford visitors or at meetings in this country or abroad. Physicians have generally been enthusiastic about the potential of these programs and what they reveal about current approaches to computer-based medical decision making.

B. Sharing and Interaction with Other SUMEX-AIM Projects

GUIDON/NEOMYCIN retains strong contact with the ONCOCIN project, as both are siblings of the MYCIN parent. These projects share programming expertise and utility routines. In addition, the central SUMEX development group acts as an important clearing house for solving problems and distributing new methods.

C. Critique of Resource Management

The SUMEX staff has been extremely helpful in maintaining connections between Xerox D-machines and SUMEX. The SUMEX group's aggressive program of converting computing resources to the workstation environment has been an outstanding success. Computing for the average programmer at any hour of the day is better than anyone experienced ten years ago, except for early morning on an unloaded system. With the addition of bit-map graphics within the programming environment and sufficient printers, we find that average preparation time for a figure to be inserted in a paper has dropped from several hours to one hour or even just a few minutes.

III. RESEARCH PLANS

A. Project Goals and Plans

Research over the next year will continue on several fronts, including one or more prototype instructional programs.

1. Use of GUIDON-DEBUG by medical students to empirically develop the interface and teaching scenario. In this setting, we will have physician teachers watch and advise as students use the program, using this opportunity to formulate their teaching strategies.

2. Incremental improvements to GUIDON-DEBUG, including work on editing the patient-specific model, evaluating student annotations of the transcript, and probably some kind of advising or coaching capability.

3. Starting in the fall, we will begin implementing a new explanation program that will produce clear text, use a student model, and be integrated with the display package.

4. A new project idea is to have online text, such as from the book Medicine, by Fishman, et al., tied to the knowledge base. A replica of the book, with fonts, will be displayed on the screen. Selecting text will display the relevant portion of the knowledge base. This will be a good entry point for students, as well as providing an interesting introduction to the idea of a knowledge base for demonstrations to physicians.

5. Another new idea which would be easy and useful to do, is to have a means for recording interactions with the program. These sessions would then be

edited for redisplay as automatic demonstrations or as introductions for new students. Text, questions, and the opportunity to try certain features will involve the viewer and test his understanding. We call this system GUIDON-SCRIPTS. This capability to replay display sequences would also be invoked automatically by explanation and teaching rules.

B. Long term plans: the GUIDON2 Family of Instructional Programs

In last year's annual report, we described our research program for 1985-88. Here, we provide a summary of the most important projects that may be completed in the next few years.

GUIDON-MANAGE In this system the student solves a problem by telling NEOMYCIN what task to do at each step. Essentially, the student provides the strategy and the program supplies the tactics (meta-rules) and domain knowledge to carry out the strategy. The program will in general carry through tasks in a logical way, for example, proceeding to test a hypothesis completely, and not "breaking" on low-level tasks that mainly test domain knowledge rather than strategy. The program will not pursue new hypotheses automatically. However, the student will always see what questions a task caused the program to request, as well as how the differential changes. This activity leads the student to observe what a strategy entails, helping him become a better observer of his own behavior. Here he shows that he knows the structural vocabulary that makes a strategy appropriate.

GUIDON-APPRENTICE This is a variant of NEOMYCIN in which the program stops during a consultation and asks the student to propose the next data request(s). The student is asked to indicate the task and focus he has in mind, plus the differential he is operating upon. The program compares this proposal to what NEOMYCIN would do. In this activity we descend to the domain level and require the student to instantiate a strategy appropriately. Ultimately, such a program will use a *learning model* that anticipates what the student is ready to learn next and how he should be challenged. Early versions can simply use built-in breakpoints supplied by an expert teacher. In the future, programs will develop their own curricula from a case library.

GUIDON-SOLVE This is the complete tutorial system. The student carries through diagnosis completely, while a student modeling program attempts to track what he is doing and a coach interrupts to offer advice. Here annotation, comparison, debugging, and explanation are all integrated to illustrate to the student how his solution is non-optimal. For example, the student might be asked to annotate his solution after he is done; this will point out strategic gaps in his awareness and provide a basis for critique and improvement. A "curriculum" based on frequent student faults and important things to learn will drive the interaction. In this activity, the student is on his own. Faced with the proverbial "blank screen," he must exercise his diagnostic procedure from start to finish.

Now that we are producing instructional programs from NEOMYCIN which can be tested, we will be emphasizing empirical work with students in all of our future research.

C. Requirements for Continued SUMEX Use

The D-machine's large address space permits development of the large programs that complex computer-aided instruction requires. Graphics enable us to develop new methods for presenting material to naive users. We use the D-machine as a reliable, constant "load-average" machine, for running experiments with physicians and students. The development of GUIDON2 on the D-machine demonstrates the feasibility of running intelligent consultation or tutoring systems on small, affordable machines in physicians' offices, schools, and other remote sites.

D. Requirements for Additional Computing Resources

We believe that one additional D-machine would improve our computing efficiency. Experience indicates that one machine per person is desirable, allowing each person to preserve his work and begin again without delay.

E. Recommendations for Future Community and Resource Development

All GUIDON and NEOMYCIN work has shifted shifting to Xerox Dolphins and Dandelions (D-machines), with the exception of daily mail and text editing, which is performed on the DEC 2060. Given that D-machines cannot be accessed remotely, work at home is restricted to what can be done on the 2060. We believe that the availability of personal workstations (the D-machines) lessens the need to do programming at home. However, we still find it convenient to be able to write at home. Integration of the workstation graphics with our wordprocessor, made available at home, would greatly increase our ability to write and prepare presentations.

Difficulties with the current resource include: unreliable file server for D-machines and inadequate utilities for examining a directory and deleting files. While the situation has greatly improved over the past few years, basic development is still required. Given the amount of software delivered in the recent past, the delay in producing an integrated environment is understandable. It is clear that a central resource group will be necessary for some time to come, charged with integrating networks, file servers, and attached devices to respond to individual requests.

E. H. Shortliffe

IV.A.2. MOLGEN Project

MOLGEN - Applications of Artificial Intelligence to Molecular
Biology: Research in Theory Formation, Testing, and Modification

Prof. E. Feigenbaum and Dr. P. Friedland
Department of Computer Science
Stanford University

Prof. Charles Yanofsky
Department of Biology
Stanford University

I.   SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

The MOLGEN project has focused on research into the applications of symbolic computation and inference to the field of molecular biology. This has taken the specific form of systems which provide assistance to the experimental scientist in various tasks, the most important of which have been the design of complex experiment plans and the analysis of nucleic acid sequences. Our current research concentrates on scientific discovery within the subdomain of regulatory genetics. We desire to explore the methodologies scientists use to modify, extend, and test theories of genetic regulation, and then emulate that process within a computational system.

Theory or model formation is a fundamental part of scientific research. Scientists both use and form such models dynamically. They are used to predict results (and therefore to suggest experiments to test the model) and also to explain experimental results. Models are extended and revised both as a result of logical conclusions from existing premises and as a result of new experimental evidence.

Theory formation is a difficult cognitive task, and one in which there is substantial scope for intelligent computational assistance. Our research is toward building a system which can form theories to explain experimental evidence, can interact with a scientist to help to suggest experiments to discriminate among competing hypotheses, and can then revise and extend the growing model based upon the results of the experiments.

The MOLGEN project has continuing computer science goals of exploring issues of knowledge representation, problem-solving, discovery, and planning within a real and complex domain. The project operates in a framework of collaboration between the Heuristic Programming Project (HPP) in the Computer Science Department and various domain experts in the departments of Biochemistry, Medicine, and Biology. It draws from the experience of several other projects in the HPP which deal with applications of artificial intelligence to medicine, organic chemistry, and engineering.

B. Medical Relevance and Collaboration

The field of molecular biology is nearing the point where the results of current research will have immediate and important application to the pharmaceutical and chemical industries. Already, clinical testing has begun with synthetic interferon and human growth hormone produced by recombinant DNA technology. Governmental reports estimate that there are more than two hundred new and established industrial firms already undertaking product development using these new genetic tools.

The programs being developed in the MOLGEN project have already proven useful and important to a considerable number of molecular biologists. Currently several dozen researchers in various laboratories at Stanford (Prof. Paul Berg's, Prof. Stanley Cohen's, Prof. Laurence Kedes', Prof. Douglas Brutlag's, Prof. Henry Kaplan's, and Prof. Douglas Wallace's) and over four hundred others throughout the country have used MOLGEN programs over the SUMEX-AIM facility. We have exported some of our programs to users outside the range of our computer network (University of Geneva [Switzerland], Imperial Cancer Research Fund [England], and European Molecular Biology Institute [Heidelberg] are examples). The pioneering work on SUMEX has led to the establishment of a separate NIH-supported facility, BIONET, to serve the academic molecular biology research community with MOLGEN-like software. BIONET is now serving many of the computational needs of over two thousand academic molecular biologists in the United States.

C. Highlights of Research Progress

C.1 Accomplishments

During the past year we have concentrated on the qualitative modeling and simulation aspects of the research. Our view is that a well-formulated, multi-level model of a scientific theory is a necessary first step to automated discovery. In addition, we have worked on knowledge acquisition and graphical display of process information and on the description and understanding of the results of laboratory experiments. The highlights of this work are summarized in several categories below.

C.1.1 Qualitative Modeling and Simulation

Our work in qualitative simulation has been directed towards building a program which embodies a theory of the tryptophan system. This simulator will contain two different types of qualitative models: a model which inter-relates the values of numeric parameters of the system via constraints, and a model which contains symbolic representations of the structure and function of objects in the system. The implementation of both these models has been partially completed.

The parametric model will be similar to existing qualitative simulation programs (such as those of deKleer, Kuipers, and Iwasaki) in that it focuses upon the relations among a small set of numeric parameters which describe a system. It will be different from these systems in that both its parameters and the relations between them will be more flexible and expressive.

Parameters will be able to take on a different types of values depending on the problem at hand and the degree of precision with which biologists understand a given mechanism, e.g., quantitative (1.23), relative (2 times Normal), or qualitative (High). Relations will be expressed with what we term "malleable constraints". These rectify an important shortcoming we have noted in previous work on qualitative simulation. The motivations behind this work are that qualitative representations should be used instead of quantitative because quantitative information is often unavailable, or can obscure the essence of a problem and make explanation more difficult. Thus, researchers have chosen to let parameters in their models take on qualitative values such as {-1, 0, 1}. We propose that these same concerns apply to the *relations among* parameters as well as their values. Just as we might know the values of different parameters with different degrees of precision, so we might understand the relationship between parameters more or less precisely. Thus we strive to represent constraints themselves in a qualitative fashion when the precise form of a relation is not known, and are studying how to propagate different types of values through these malleable constraints.

We have begun to design a language for expressing constraints with varying degrees of precision. A very imprecise constraint would state only that two parameters vary

directly (or inversely). Adding more precision could tell us the form of the constraint, e.g., is it linear, quadratic, or exponential? More precision still would tell us the value of the constant for a linear constraint.

We have constructed a knowledge base which describes roughly 15 parameters and the relations between them, and we are currently able to propagate qualitative values through this network. Work has begun on representing malleable constraints and other types of values.

The structural model should describe the mechanisms of the trp system at several different levels of detail. This has benefits for both problem solving and discovery.

Processes and objects can be represented at different levels of detail by abstracting away elements of their description. For example, the process of transcription can be thought of as simply consisting of subprocesses of initiation, elongation, and termination, yet each of these subprocesses is known in greater detail. Likewise, the tryptophan operon can be thought of abstractly as a sequence of DNA segments (such as genes), or as a long, detailed list of specific nucleotides. Simpler descriptions can be reasoned about more quickly and will give reliable predictions for certain questions. They also provide an abstraction hierarchy for the discovery process, allowing a discovery module to focus on increasingly detailed sub-parts of a model to efficiently determine where the error in a theory lies.

C.1.2 Process Description and Graphical Display

A system has been built which generalizes our experience in process description by providing a simplified interface for the domain-independent description and animation of process knowledge. The system allows processes to be broken down into component sub-processes and the causal and time-oriented relationships of the subprocesses to be specified. In addition, objects utilized by the processes can be conveniently described and "drawn" with modes and points of interaction among the objects given by the user. All knowledge about processes and objects is automatically stored in the framework of a KEE knowledge base.

After process and object description, the system automatically animates the process by displaying one of several primitive types of interactions among objects in the proper time order dictated by the process knowledge base. This system has been tested on the tryptophan operon domain and its utility is currently being explored in a medical simulation domain.

C.1.3 Acquisition of Experimental Data

The KEE ActiveImages facility has been used to quickly build a convenient interface for biologists to describe the results of genetic regulation experiments. Information is entered by touching (with a mouse-operated cursor) one of several graphical gauges representing such information as absolute and relative cell growth rates. The values provided are automatically checked for consistency and entered in the frame-based experiment description knowledge base. Quantitative values are translated to qualitative values for use by the modeling and simulation system.

C.2 Research in Progress

Over the next several months we will continue our work in qualitative simulation, modeling, and process description. The full qualitative-quantitative parameter model will be completed and tested over a wide range of possible values in the trp operon system. The first version of the imprecise constraint description language will be completed. Further experiments will be performed on reasoning about processes in a hierarchy of abstraction space. The generic process description and animation system will be tested in a different, most likely medical, domain.

D. *Publications*

1. Bach, R., Friedland, P., Brutlag, D., and Kedes, L.: *MAXIMIZE, a DNA sequencing strategy advisor.* Nucleic Acids Res. 10(1):295-304, January, 1982

2. Bach, R., Friedland, P., and Iwasaki, Y.: *Intelligent computational assistance for experiment design.* Nucleic Acids Res. 12(1):11-29, January, 1984.

3. Brutlag, D., Clayton, J., Friedland, P. and Kedes, L.: *SEQ: A nucleotide sequence analysis and recombination system.* Nucleic Acids Res. 10(1):279-294, January, 1982.

4. Clayton, J. and Kedes, L.: *GEL, a DNA sequencing project management system.* Nucleic Acids Res. 10(1):305-321, January, 1982.

5. Feitelson, J. and Stefik, M.J.: *A case study of the reasoning in a genetics experiment.* Heuristic Programming Project Report HPP-77-18 (working paper), May, 1977.

6. Friedland, P.: *Knowledge-based experiment design in molecular genetics.* Proc. Sixth IJCAI, August, 1979, pp. 285-287.

7. Friedland P.: *Knowledge-based experiment design in molecular genetics.* Stanford Computer Science Report STAN-CS-79-760 (Ph.D. thesis), December, 1979.

8. Friedland, P., Kedes, L. and Brutlag D.: *MOLGEN--Applications of symbolic computation and artificial intelligence to molecular biology.* Proc. Battelle Conference on Genetic Engineering, April, 1981.

9. Friedland, P.: *Acquisition of procedural knowledge from domain experts.* Proc. Seventh IJCAI, August, 1981, pp. 856-861.

10. Friedland, P., Kedes, L., Brutlag, D., Iwasaki, Y. and Bach R.: *GENESIS, a knowledge-based genetic engineering simulation system for representation of genetic data and experiment planning.* Nucleic Acids Res. 10(1):323-340, January, 1982.

11. Friedland, P., and Kedes, L.: *Discovering the secrets of DNA.* Communications of the ACM, 28(11):1164-1186, November, 1985, and IEEE/Computer, 18(11):49:69, November, 1985.

12. Friedland, P. and Iwasaki Y.: *The concept and implementation of skeletal plans.* Journal of Automated Reasoning, 1(2): 161-208, 1985.

13. Friedland, P., Armstrong, P., and Kehler, T.: *The role of computers in biotechnology.* BIO\TECHNOLOGY 565-575, September, 1983.

14. Iwasaki, Y. and Friedland, P.: *SPEX: A second-generation experiment design system.* Proc. of Second National Conference on Artificial Intelligence, August, 1982, pp. 341-344.

15. Martin, N., Friedland, P., King, J. and Stefik M.J.: *Knowledge base management for experiment planning in molecular genetics.* Proc. Fifth IJCAI, August, 1977, pp. 882-887.

16. Meyers, S. and Friedland, P.: *Knowledge-based simulation of regulatory*

                                          E. H. Shortliffe

genetics in bacteriophage Lambda. Nucleic Acids Res. 12(1):1-9, January, 1984.

17. Stefik, M. and Friedland, P.: *Machine inference for molecular genetics: Methods and applications.* Proc. of NCC, June, 1978.

18. Stefik, M.J. and Martin N.: *A review of knowledge based problem solving as a basis for a genetics experiment designing system.* Stanford Computer Science Report STAN-CS-77-596, March, 1977.

19. Stefik, M.: *Inferring DNA structures from segmentation data: A case study.* Artificial Intelligence 11:85-114, December, 1977.

20. Stefik, M.: *An examination of a frame-structured representation system.* Proc. Sixth IJCAI, August, 1979, pp. 844-852.

21. Stefik, M.: *Planning with constraints.* Stanford Computer Science Report STAN-CS-80-784 (Ph.D. thesis), March, 1980.

E. Funding Support

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

SUMEX-AIM continues to provide the bulk of our computing resources. The facility has not only provided excellent support for our programming efforts but has served as a major communication link among members of the project. Systems available on SUMEX-AIM such as INTERLISP, TV-EDIT, and BULLETIN BOARD have made possible the project's programming, documentation and communication efforts. The interactive environment of the facility is especially important in this type of project development.

We strongly approve of the network-oriented approach to a programming environment into which SUMEX has evolved. The ability to utilize LISP workstations for intensive computing while still communicate with all of the other SUMEX resources has been very valuable to our work. We currently have a satisfactory mode of operation where most programming takes place on the workstations and most electronic communications, information sharing, and document preparation takes place within the mature TOPS-20 environment. The evolution of SUMEX has alleviated most of our previous problems with resource loading and file space. Our current workstations are not quite fast nor sophisticated enough, but we are encouraged by the progress that has been made.

We have taken advantage of the collective expertise on medically-oriented knowledge-based systems of the other SUMEX-AIM projects. In addition to especially close ties with other projects at Stanford, we have greatly benefited by interaction with other projects at yearly meetings and through exchange of working papers and ideas over the system.

The ability for instant communication with a large number of experts in this field has been a determining factor in the success of the MOLGEN project. It has made possible the near-instantaneous dissemination of MOLGEN systems to a host of experimental users in laboratories across the country. The wide-ranging input from these users has greatly improved the general utility of our project.

We find it very difficult to find fault with any aspect of the SUMEX resource management. It has made it easy for us to expand our user group, to give demonstrations (through the 20/20 adjunct system as well as the LISP workstations), and to disseminate software to non-SUMEX users overseas.

III. RESEARCH PLANS

A. Project Goals And Plans

Our current work has the following major goals:

1. Use the knowledge base to explain observations that are indeed explainable without changes to the current model. For example, "I have observed a mutation that causes constitutive (uncontrolled) production of tryptophan. How can that be explained within the Jacob-Monod model?" This process will be accomplished by some combination of forward simulation and backward rule-chaining.

2. Begin to recognize when observations are "interesting." Interesting here has one of the following broad meanings:

   a. A seeming direct contradiction to the existing theory.

   b. A statistically rare occurrence (one that is understandable by the current theory, but should not occur very often).

   c. A dramatic confirmation of the existing model.

   d. An observation currently unpredictable by the current model because the model is either not detailed enough or incomplete. The observation in this case must have a relation to the model because an important object of the model is involved or it relates to an effect predicted by the model.

3. Build a mechanism for postulating extensions or corrections to the current theory: a contrained regulatory theory generator. The overall approach to this mechanism is perhaps the most interesting problem in our work. In discussions with other computer scientists, the notion of "or" reasoning where the theory construction process consists of hierarchical refinement of abstract ideas into more detailed ones, and "and" reasoning where the theory is built up in little pieces at many different levels simultaneously has emerged. We see strong evidence for both types of reasoning within Yanofsky's project. In fact, as stated above, the global model of Yanofsky's laboratory is a hybrid one. Individual graduate students performed "and" tasks--filling in details of seemingly unrelated pieces of the model. Yanofsky was the master "or" reasoner, slowly building a hierarchical model of the new regulatory mechanism. It is in this area of our research where the greatest discussion with AI colleagues is needed and which may produce the most significant AI benefits.

4. Build a mechanism for evaluating alternative theories. This would include rating the theories based on plausibility, selectability, completeness, significance, and so on. We hope the evaluation process produces information useful in discriminating among the possible theories.

5. Test the entire structure on the evolving trp operon regulatory system. Experiment with different initial knowledge bases to see how the discovery

process is altered by the availability of new techniques, analogous systems, and so forth.

B. Justification and Requirements for Continued SUMEX Use

The MOLGEN project depends heavily on the SUMEX facility. We have already developed several useful tools on the facility and are continuing research toward applying the methods of artificial intelligence to the field of molecular biology. The community of potential users is growing nearly exponentially as researchers from most of the biomedical-medical fields become interested in the technology of recombinant DNA. We believe the MOLGEN work is already important to this growing community and will continue to be important. The evidence for this is an already large list of pilot exo-MOLGEN users on SUMEX.

We support with great enthusiasm the acquisition of satellite computers for technology transfer and hope that the SUMEX staff continues to develop and support these systems. One of the oft-mentioned problems of artificial intelligence research is exactly the problem of taking prototypical systems and applying them to real problems. SUMEX gives the MOLGEN project a chance to conquer that problem and potentially supply scientific computing resources to a national audience of biomedical-medical research scientists.

Responses to Questions Regarding Resource Future

1. Role of SUMEX after 7/86--I strongly believe that the 2060 should have continuing support for the foreseeable future. The maturity of software for communications, document preparation, and general support of scientific literacy is unsurpassed. One has only to note the heavy continued load on SUMEX, despite the proliferation of workstations, VAXes, etc. around the KSL to see that it is still being used productively. In addition, the ability to easily work from home at all hours contributes greatly to overall productivity within the SUMEX community.

2. Will my group require continued access--Yes, very much so for all of the reasons outlined above.

3. Impact of user fees--Modest user fees would not have an enormous impact, but would prevent the kind of easy, productive use for general purposes that SUMEX now serves. I think the greater impact would be on not fully established or new research groups during start-up mode.

4. Workstation plans--My group, MOLGEN, already makes extensive use of workstations for mainline computing purposes. Despite this use, we still find the SUMEX 2060 invaluable.

I would add to #1, that continuing research on melding together a distributed environment (of which both single-user workstations and the 2060 are parts) should be a major continuing goal of SUMEX research.

IV.A.3. ONCOCIN Project

ONCOCIN Project

Edward H. Shortliffe, M.D., Ph.D.
Departments of Medicine and Computer Science
Stanford University

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

The ONCOCIN Project is one of many Stanford research programs devoted to the development of knowledge-based expert systems for application to medicine and the allied sciences. The central issue in this work has been to develop a program that can provide advice similar in quality to that given by human experts, and to ensure that the system is easy to use and acceptable to physicians. The work seeks to improve the interactive process, both for the developer of a knowledge-based system, and for the intended end user. In addition, we have emphasized clinical implementation of the developing tool so that we can ascertain the effectiveness of the program's interactive capabilities when it is used by physicians who are caring for patients and are uninvolved in the computer-based research activity.

B. Medical Relevance and Collaboration

The lessons learned in building prior production rule systems have allowed us to create a large oncology protocol management system much more rapidly than was the case when we started to build MYCIN. We introduced ONCOCIN for use by Stanford oncologists in May 1981. This would not have been possible without the active collaboration of Stanford oncologists who helped with the construction of the knowledge base and also kept project computer scientists aware of the psychological and logistical issues related to the operation of a busy outpatient clinic.

C. Highlights of Research Progress

C.1 Background and Overview of Accomplishments

The ONCOCIN Project is a large interdisciplinary effort that has involved over 35 individuals since the project's inception in July 1979. The work is currently in its seventh year; we summarize here the milestones that have occurred in the research to date:

- *Year 1:* The project began with two programmers (Carli Scott and Miriam Bischoff), a Clinical Specialist (Dr. Bruce Campbell) and students under the direction of Dr. Shortliffe and Dr. Charlotte Jacobs from the Division of Oncology. During the first year of this research (1979-1980), we developed a prototype of the ONCOCIN consultation system, drawing from programs and capabilities developed for the EMYCIN system-building project. During that year, we also undertook a detailed analysis of the day-to-day activities of the Stanford Oncology Clinic in order to determine how to introduce ONCOCIN with minimal disruption of an operation which is already running smoothly. We also spent much of our time in the first year giving careful consideration to the most appropriate mode of interaction with physicians in order to optimize the chances for ONCOCIN to become a useful and accepted tool in this specialized clinical environment.

- *Year 2:* The following year (1980-1981) we completed the development of a special interface program that responds to commands from a customized keypad. We also encoded the rules for one more chemotherapy protocol (oat cell carcinoma of the lung) and updated the Hodgkin's Disease protocols when new versions of the documents were released late in 1980; these exercises demonstrated the generality and flexibility of the representation scheme we had devised. Software protocols were developed for achieving communication between the interface program and the reasoning program, and we coordinated the printing routines needed to produce hard copy flow sheets, patient summaries, and encounter sheets. Finally, lines were installed in the Stanford Oncology Day Care Center, and, beginning in May 1981, eight fellows in oncology began using the system three mornings per week for management of their patients enrolled in lymphoma chemotherapy protocols.

- *Year 3:* During our third year (1981 - 1982) the results of our early experience with physician users guided both our basic and applied work. We designed and began to collect data for three formal studies to evaluate the impact of ONCOCIN in the clinic. This latter task required special software development to generate special flow sheets and to maintain the records needed for the data analysis. Towards the end of 1982 we also began new research into a *critiquing model* for ONCOCIN that involves "hypothesis assessment" rather than formal advice giving. Finally, in 1982 we began to develop a query system to allow system builders as well as end users to examine the growing complex knowledge base of the program.

- *Year 4:* Our fourth year (1982-1983) saw the departure of Carli Scott, a key figure in the initial design and implementation of ONCOCIN, the promotion of Miriam Bischoff to Chief Programmer, and the arrival of Christopher Lane as our second scientific programmer. At this time we began exploring the possibility of running ONCOCIN on a single-user professional workstation and experimented with different options for data-entry using a "mouse" pointing device. Christopher Lane became an expert on the Xerox workstations that we are using. In addition, since ONCOCIN had grown to such a large program with many different facets, we spent much of our fourth year documenting the system. During that year we also modified the clinic system based upon feedback from the physician-users, made some modifications to the rules for Hodgkin's disease based upon changes to the protocols, and completed several evaluation studies.

- *Year 5:* The project's fifth year (1983-1984) was characterized by growth in the size of our staff (three new full-time staff members and a new oncologist joined the group). The increased size resulted from a DRR grant that permitted us to begin a major effort to rewrite ONCOCIN to run on professional workstations. Dr. Robert Carlson, who had been our Clinical Specialist for the previous two years, was replaced by Dr. Joel Bernstein, while Dr. Carlson assumed a position with the nearby Northern California Oncology Group; this appointment permitted him to continue his affiliation both with Stanford and with our research group. In August of 1983, Larry Fagan joined the project to take over the duties of the ONCOCIN Project Director while also becoming the Co-Director of the newly formed Medical Information Sciences Program. Dr. Fagan continues to be in charge of the day-to-day efforts of our research. An additional programmer, Jay Ferguson, joined the group in the fall to assist with the effort required to transfer ONCOCIN from SUMEX to the 1108 workstation. A fourth programmer, Joan Differding, joined the staff to work on our protocol acquisition effort (OPAL).

E. H. Shortliffe   110

- *Year 6:* During our sixth year (1984-1985) we have further increased the size of our programming staff to help in the major workstation conversion effort. The ONCOCIN and OPAL efforts were greatly facilitated by a successful application for an equipment grant from Xerox Corporation. With a total of 15 Xerox LISP machines now available for our group's research, all full time programmers have dedicated machines, as do several of the senior graduate students working on the project. Christopher Lane took on full-time responsibility for the integration and maintenance of the group's equipment and associated software. Two of our programming staff moved on to jobs in industry (Bischoff and Ferguson) and three new programmers (David Combs, Cliff Wulfman, and Samson Tu) were hired to fill the void created by their departure and by the reassignment of Christopher Lane.

In addition to funding from DRR for the workstation conversion effort, we have support from the National Library of Medicine which supports our more basic research activities regarding biomedical knowledge representation, knowledge acquisition, therapy planning, and explanation as it relates to the ONCOCIN task domain. We have continued to study the therapy planning process under support from the NLM. This research is led by Dr. Fagan and has concentrated on how to represent the therapy-planning strategies used to decide treatment for patients who run into serious problems while on protocol-described treatment. The physicians who treat these patients often seek out a consultation with the protocol study chairman. Dr. Branimar Sikic, a faculty member from the Stanford University Department of Medicine, and the Study Chairman for the oat cell protocol, is collaborating on this project. Janice Rohn joined the ONCOCIN project as data manager and to assist in the knowledge entry process.

- *Year 7:* This year (1985-86) marked several milestones in our research on workstation-based programming. The OPAL knowledge acquisition system became operational, and several new oncology protocols have been entered using this system. David Combs has been primarily responsible for creating the operational version of OPAL (based on the initial prototype by Joan Differding Walton). As anticipated, we have increased the speed and ease with which protocols can be added to the ONCOCIN knowledge base.

Based on the protocols entered through OPAL, we have begun experimental testing of the workstation version of ONCOCIN in the Stanford oncology clinic. Clifford Wulfman developed the user interface (based on an initial prototype designed by Christopher Lane). Samson Tu developed the reasoning component (designed originally by Jay Ferguson). Much of their work is built upon an object-oriented system developed for our group by Christopher Lane. We have connected the various parts of the system, and have demonstrated that we have the capability to run ONCOCIN with the reasoning program and interface program on different machines in the communication network. The current version of the program is currently run on a single workstation, but future versions may take advantage of the multiple machine option. To increase the speed at which we are able to test protocols entered into ONCOCIN, we have developed additional programs to test real and synthetic cases without user interaction; these are then reviewed by our collaborating clinicians.

We have also developed a workstation-based program, OPUS, to help clinicians determine which protocols are appropriate for specific patients. OPUS was designed and implemented by Janice Rohn with the assistance of

Christopher Lane. We have been using it in the clinic setting since the end of 1985. Thus, in addition to providing an information resource about protocols, the use of a graphically-oriented program provided a way to learn about the software style and hardware used in the workstation version of ONCOCIN.

We discontinued the mainframe version of ONCOCIN, and are using the workstation version exclusively. The performance of the mainframe version of ONCOCIN was documented in two evaluation papers that appeared in clinical journals (see Hickam and Kent papers).

We have continued our basic research in the design of advanced therapy-planning programs: the ONYX project. We have developed a model for planning which includes techniques from the fields of artificial intelligence, simulation, and decision analysis. Artificial intelligence techniques are used to create a small number of possible plans given the ideal therapy and the patient's past treatment history. Simulation techniques and decision analysis are used to examine and order the most promising plans. Our goal is to allow ONCOCIN to give advice in a wider range of situations; in particular, the system should be able to recommend plans for patients who have an unusual response to chemotherapy.

During this year, Stephen Rappaport, M.D. joined us as a programmer on the therapy planning research. Clinical expertise for ONCOCIN was provided by Richard Lenon, M.D. and Robert Carlson, M.D.

C.2 Research in Progress

Our research in the ONCOCIN project over the last year comprised three major categories: (1) conversion of ONCOCIN to the workstation version, (2) development of a knowledge acquisition interface (OPAL) for entering new protocols, and (3) modeling of the strategic therapy selection process (ONYX). We are now able to explore ways to test the system beyond the Stanford environment.

A summary of our current research endeavors follows.

C.2.1 Transfer of the ONCOCIN system from the DEC-20 to the Xerox 1100 Series machines

During the process of converting to the workstation version of ONCOCIN, we redesigned segments of the program. We have completed the major portion of that work, and our experience with the new version has suggested additional areas for improving the reasoning techniques and knowledge representation of ONCOCIN.

- *Redesign of the reasoning component.* A major impetus for the redesign of the system was to develop more efficient methods to search the knowledge base during the running of a case. We have implemented a reasoning program that uses a discrimination network to process the cancer protocols. This network provides for a compact representation of information which is common to many protocols but does not require the program to consider and then disregard information related to protocols that are irrelevant to a particular patient. We continue to improve portions of the reasoning component that are associated with reasoning over time; e.g., modeling the appropriate timing for ordering tests and identifying the information which needs to be gathered before the next clinic visit. In general, we are concentrating on improving the representation of the knowledge regarding sequences of therapy actions specified by the protocol.

We are also improving the reasoning component's efficiency, such as by building concurrency into the reasoning process. Our overall aim in this area is to increase generality of the reasoning program; our long-term goal is to develop E-ONCOCIN, a domain-independent, time-oriented reasoning program.

- *Development of a temporal network.* The ability to represent temporal information is a key element of programs that must reason about treatment protocols. The earlier version of the ONCOCIN system did not have an explicit structure for reasoning about time-oriented events. We are experimenting with different configurations of the temporal network, and with the syntax for querying the network. We are also adapting this network so that it can interface with the ONYX therapy-planning systems.

- *Extensions to the user interface.* We continue to experiment with various configurations of the user interface, including new formats for reviewing the therapy recommendations and specifying tests to be ordered. We have also developed a program which allows the data manager and end users to start up their interaction with ONCOCIN.

  A continuing area of research concerns how to guide the user to the most appropriate items to enter (based on the needs of the reasoning program) without disrupting the fixed layout of the flowsheet. The mainframe version of ONCOCIN modified the flowsheet in order to extract necessary information from the user. In the workstation version, we have developed a guidance mechanism which alerts the user to items that are needed by the reasoning program. The user is not required to deviate from a preferred order of entry nor required to respond to a question for which no current answer is available.

- *System support for the reorganization.* The LISP language, which we used to build the first version of ONCOCIN, does not explicitly support basic knowledge manipulation techniques (such as message passing, inheritance techniques, or other object-oriented programming structures). These facilities are available in some commercial products, but none of the existing commercial implementations provide the reliability, speed, size, or special memory-manipulation techniques that are needed for our project. We have therefore developed a "minimal" object-oriented system to meet our specifications. The object system is currently in use by each component of the new version of ONCOCIN and in the software used to connect these components. In addition, all ONCOCIN student projects are now based on this programming environment.

C.2.2 Interactive Entry of Chemotherapy Protocols by Oncologists (OPAL)

A major effort in this grant year has been the development and testing of software (the OPAL system) that will permit physicians who are not computer programmers to enter protocol information on a structured set of forms presented on a graphics display. Most expert systems require tedious entry of the system's knowledge. In many other medical expert systems, each segment of knowledge is transferred from the physician to the programmer, who then enters the knowledge into the expert system. We have taken advantage of the generally well-structured nature of cancer treatment plans to design a knowledge entry program that can be used directly by clinicians. The structure of cancer treatment plans includes:

- choosing among multiple protocols (that may be related to each other);

- describing experimental research arms in each protocol;

- specifying individual drugs and drug combinations;

- setting the drug dosage level;

- and modifying either the choice of drugs or their dosage.

Using the graphics-oriented workstations, this information is presented to the user as computer-generated forms which appear on the screen. After the user fills in the blanks on the forms, the program generates the rules used to drive the reasoning process. As the user describes more detailed aspects of the protocol, new forms are added to the computer display; these allow the user to specify the special cases that make the protocols so complicated. Although the user is unaware of the creation of the knowledge base from the interaction with OPAL, a complex set of translations are taking place. The user's entries are mapped into an intermediate data structure (IDS) that is common for all protocols. From the IDS, a translation program generates rules for creating and modifying treatment, and integrates them with the existing ONCOCIN knowledge base. Improving the design of the IDS and the rule translation programs will be a major research effort of this year.

Although the "forms" were specifically designed for cancer treatment plans, the techniques used to organize data can be extended to other clinical trials, and eventually to other structured decision tasks. The key factor is to exploit the regularities in the structure of the task (e.g., this interface has an extensive notion of how chemotherapy regimens are constructed) rather than to try to build a knowledge-entry system that can accept *any* possible problem specification. The OPAL program is based upon a domain-independent forms creation package designed and implemented by David Combs. This program will provide the basis for our extension of OPAL to other application areas.

We have entered six protocols covering many different organ systems and styles of protocol design. Based on this experience, we are modifying OPAL to increase the percentage of the protocol that can be entered directly by our clinical collaborators. One direction in which we have extended the OPAL program is in providing a graphical interface of nodes and arcs to specify the procedural knowledge about the order of treatments and important decision points within the treatments. This work is described in several papers by Musen.

C.2.3 Strategic Therapy Planning (ONYX)

As mentioned above, we have continued our research project (ONYX) to study the therapy-planning process and to determine how clinical strategies are used to plan therapy in unusual situations. Our goals for ONYX are: (1) to conduct basic research into the possible representations of the therapy-planning process, (2) to develop a computer program to represent this process, and (3) eventually to interface the planning program with ONCOCIN. The project members (Fagan, Kahn, Langlotz, Rappaport, and Tu) have spent many hours meeting with Dr. Sikic to determine how he plans therapy for patients whose special clinical situation precludes following the standard therapeutic plan described in the protocol document.

The prototype program design has four components: (1) to review the patient's past record and recognize emerging problems, (2) to formulate a small number of revised therapy plans based on existing problems, (3) to determine the results of the generated plans by using simulation, and (4) to weight the results of the simulation and rank

order the plans by performing decision analysis. This model is described in the papers by Langlotz.

C.2.4 Documentation

We recently videotaped a lecture and demonstration of the ONCOCIN and OPAL systems at the XEROX Palo Alto Research Center. This videotape will be available for loan from our offices. Our previous videotapes have been shown at scientific meetings and have been distributed to many researchers in other countries. The publications described below further document our recent work on ONCOCIN.

C.2.5 Dissemination

We are planning experimental installation of ONCOCIN workstations in private oncology offices in San Jose and San Francisco. An application proposing this project is currently under review.

D. Publications Since January, 1985

1. Hickam, D.H., Shortliffe, E.H., Bischoff, M.B., Scott, A.C., Jacobs, C.D. A study of the treatment advice of a computer-based cancer chemotherapy protocol advisor (Memo KSL-85-21). Annals of Internal Medicine 103(6 pt 1):928-936 (1985).

2. Kent, D.L., Shortliffe, E.H., Carlson, R.W., Bischoff, M.B., Jacobs, C.D. Improvements in data collection through physician use of a computer-based chemotherapy treatment consultant (Memo KSL-85-22). Journal of Clinical Oncology 3:1409-1417 (1985).

3. Tsuji, S. and Shortliffe, E.H. Graphical access to a medical expert system: I. Design of a knowledge engineer's interface (Memo KSL-85-11). Meth. Inf. Med., Vol. 25, April 1986.

4. Preston, K.,Jr., Fagan, L.M., Huang, H.K., Pryor, T.A. Computing in medicine. IEEE Computer 17(10):294-313, October 1984.

5. Langlotz, C., Fagan, L., Tu, S., Williams, J., Sikic, B. ONYX: An architecture for planning in uncertain environments (Memo KSL-85-10). Proceedings of the Ninth International Joint Conference on Artificial Intelligence, pp.447-449, Los Angeles, CA, August 1985.

6. Lane, C.D., Differding, J.C., Shortliffe, E.H. Graphical access to a medical expert system: II. Design of an interface for physicians (Memo KSL-85-15). To appear in Meth. Inf. Med., July 1986.

7. Musen, M., Langlotz, C., Fagan, L., Shortliffe, E.H. Rationale for knowledge base redesign in a medical advice system (Memo KSL-85-17). Proceedings of AAMSI Congress 85 (A. Levy and B. Williams, Eds.), pp. 197-201, San Francisco, May 20-22, 1985.

8. Fagan, L. New directions for expert systems: Examples from the ONCOCIN project. Proceedings of AAMSI Congress 85 (A. Levy and B. Williams, Eds.), pp. 183-186, San Francisco, May 20-22, 1985.

9. Kahn, M.G., Ferguson, J. C., Shortliffe, E.H., Fagan, L. Representation and Use of Temporal Information in ONCOCIN. (Memo KSL-85-8). Proceedings of the 9th SCAMC, pp. 172-176, Baltimore, MD, November 1985.

10. Fagan, L., Differding, J., Langlotz, C., Tu, S. Knowledge acquisition and strategic therapy planning for cancer clinical trials. Proceedings of the International Conference on Artificial Intelligence in Medicine (I. De Lotto and M. Stefanelli, Eds.), Pavia, Italy, 13-14 September 1985.

11. Musen, M.A., Rohn, J.A., Fagan, L.M., Shortliffe, E.H. Knowledge engineering for a clinical trial advice system: uncovering errors in protocol specification (Memo KSL-85-51). Proceedings of AAMSI Congress 86, Anaheim, CA, May 8-10, 1986.

12. Langlotz, C.P., Fagan, L.M., Shortliffe, E.H. Overcoming limitations of artificial intelligence planning techniques (Memo KSL-85-52). Proceedings of AAMSI Congress 86, Anaheim, CA, May 8-10, 1986.

13. Musen, M.A., Fagan, L.M., Shortliffe, E.H. Graphical specification of procedural knowledge for an expert system (Memo KSL-85-53). December 1985. To be presented at the Second IEEE Computer Society Workshop on Visual Languages, Dallas, TX, June 1986.

14. Combs,D.M., Musen, M.A., Fagan, L.M., Shortliffe, E.H. Graphical entry of procedural and inferential knowledge (Memo KSL-85-56). Proceedings of AAMSI Congress 86, Anaheim, CA, May 8-10, 1986.

15. Lane, C.D., Frisse, M.E., Fagan, L.M., and Shortliffe, E.H. Object-oriented graphics in medical interface design (Memo KSL-85-58). Proceedings of AAMSI Congress 86, Anaheim, CA, May 8-10, 1986.

16. Musen, M.A., Fagan, L.M., Combs, D.M., Shortliffe, E.H. Facilitating knowledge entry for an oncology therapy advisor using a model of the application area (Memo KSL-86-1). To appear in Proceedings of MEDINFO-86, October 1986.

17. Langlotz, C.P., Fagan, L.M., Tu, S.W., Sikic, B.I., Shortliffe, E.H. Combining artificial intelligence and decision analysis for automated therapy planning assistance (Memo KSL-86-3). To appear in Proceedings of MEDINFO-86, October 1986.

18. Kahn, M.G., Fagan, L.M., Shortliffe, E.H. Context-specific interpretation of patient records for a therapy advice system (Memo KSL-86-4). To appear in Proceedings of MEDINFO-86, October 1986.

19. Langlotz, C.P., Shortliffe, E.H., Fagan, L.M. Using decision theory to justify heuristics (Memo KSL-86-26). To appear in Proceedings of AAAI-86, August 1986.

20. Shortliffe, E.H. Artificial Intelligence in Management Decisions: ONCOCIN (Memo KSL-86-39). Proceedings of a Conference on Medical Information Sciences, University of Texas Health Sciences Center at San Antonio, July 1985.

E. Funding Support

Agency: Biotechnology Resources Program, Division of Research Resources
ID Number: RR 01613
Term: July 1983 to June 1986
Total award: $624,455
Current award: (7/85-6/86): $200,302

Grant Title: "Therapy-planning strategies for consultation by computer"
Principal Investigator: Edward H. Shortliffe
Agency: National Library of Medicine
ID Number: LM-04136
Term: August 1983 to July 1986
Total award: $211,851
Current award: (8/85-7/86) $74,150

Grant Title: "Knowledge Management for Clinical Trial Advice Systems"
Principal Investigator: Edward H. Shortliffe
Agency: National Library of Medicine
ID Number: 1 R01 LM04420-01
Term: September 1985 through August 1988
Total award: $314,707
Current Award: (9/85-8/86): $95,205

Grant Title:"Information Structure and Use in Knowledge-based Expert Systems"
Principal Investigator: Bruce G. Buchanan
Co-Principal Investigator: Edward H. Shortliffe
Agency: National Science Foundation
ID Number: IST 83-12148
Term: March 1, 1984 - February 28, 1987
Total award: $330,000 (includes indirects)
Current award (3/85-2/86) $52,679   (Shortliffe portion)

Grant Title: Postdoctoral Training in Medical Information Science
Principal Investigator: Edward H. Shortliffe
Agency: National Library of Medicine
ID Number: 1 T32 LM07033
Term: July 1, 1984 - June 30, 1989
Total award: $903,718
Current award (7/1/85-6/30/86): $215,850

Grant Title: Henry J. Kaiser Faculty Scholar in General Internal Medicine
Principal Investigator: Edward H. Shortliffe
Agency: Henry J. Kaiser Family Foundation
Term: July 1983 to June 1988
Total award: $250,000 ($50,000 annually).

Grant Title: "Explanation of Computer-Assisted Therapy Plans"
Principal Investigator: Lawrence M. Fagan
Agency: NIH
ID Number: LM04316
Term: Feb. 1985 to Jan. 1988
Total award: $107,441.
Current award: (2/85-1/86) $37,500

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaborations and Program Dissemination via SUMEX

A great deal of interest in ONCOCIN has been shown by the medical, computer science, and lay communities. We are frequently asked to demonstrate the program to Stanford visitors (both the prototype system running in the clinic and the newer work transferring the system to professional workstations). We also demonstrated our developing workstation code in the Xerox exhibit in the trade show associated with AAAI-84 in Austin, Texas and IJCAI-85 in Los Angeles. Physicians have generally been enthusiastic about ONCOCIN's potential. The interest of the lay community is reflected in the frequent requests for magazine interviews and television coverage of the work. Articles about MYCIN and ONCOCIN have appeared in such diverse publications as *Time* and *Fortune*, and ONCOCIN has been featured on the "NBC Nightly News," the PBS "Health Notes" series, and "The MacNeil-Lehrer Report." Due to the frequent requests for ONCOCIN demonstrations, we have produced a videotape about the ONCOCIN research which includes demonstrations of our professional workstation research projects and the 2020-based clinic system. The tape has been shown at several national meetings, including the 1984 Workshop on Artificial Intelligence in Medicine, the 1984 meeting of the Society for Medical Decision Making, and the 1985 meeting of the Society for Research and Education in Primary Care Internal Medicine. The tape has also been shown to both national and international researchers in biomedical computing. We have also completed an updated tape of our activities for demonstration purposes.

Our group also continues to oversee the MYCIN program (not an active research project since 1978) and the EMYCIN program. Both systems continue to be in demand as demonstrations of expert systems technology. MYCIN has been demonstrated via networks at both national and international meetings in the past, and several medical school and computer science teachers continue to use the program in their computer science or medical computing courses. Researchers who visit our laboratory often begin their introduction by experimenting with the MYCIN/EMYCIN systems. We also have made the MYCIN program available to researchers around the world who access SUMEX using the GUEST account. EMYCIN has been made available to interested researchers developing expert systems who access SUMEX via the CONSULT account. One such consultation system for psychopharmacological treatment of depression, called Blue-Box (developed by two French medical students, Benoit Mulsant and David Servan-Schreiber), was reported in July of 1983 in *Computers and Biomedical Research.*

B. Sharing and Interaction with Other SUMEX-AIM Projects

The community created on the SUMEX resource has other benefits which go beyond actual shared computing. Because we are able to experiment with other developing systems, such as INTERNIST/CADUCEUS, and because we frequently interact with other workers (at AIM Workshops or at other meetings), many of us have found the scientific exchange and stimulation to be heightened. Several of us have visited workers at other sites, sometimes for extended periods, in order to pursue further issues which have arisen through SUMEX- or workshop-based interactions. In this regard, the ability to exchange messages with other workers, both on SUMEX and at other sites, has been crucial to rapid and efficient dissemination of ideas. Certainly it is unusual for a small community of researchers with similar scholarly interests to have at their disposal such powerful and efficient communication mechanisms, even among those researchers on opposite coasts of the country.

During this past year, we have had extensive interactions with Randy Miller at

Pittsburgh. Via floppy disks and SUMEX, we have experimented with several versions of the QUIK program. The interaction was very much facilitated by the availability of SUMEX for communication and data transmission.

C. Critique of Resource Management

Our community of researchers has been extremely fortunate to work on a facility that has continued to maintain the high standards that we have praised in the past. The staff members are always helpful and friendly, and work as diligently to please the SUMEX community as to please themselves. As a result, the computer is as accessible and easy-to-use as they can make it. More importantly, it is a reliable and convenient research tool. We extend special thanks to Tom Rindfleisch for maintaining such high professional standards. As our computing needs grow, we have increased our dependence on special SUMEX skills such as networking and communication protocols.

III. RESEARCH PLANS

A. Project Goals and Plans

In the coming year, there are several areas in which we expect to expend our efforts on the ONCOCIN System:

1. *Development of a workstation model for cost-effective dissemination of clinical consultation systems.* To meet this specific aim we will continue the basic and applied programming efforts (ONCOCIN, OPAL, and ONYX) described earlier in this report.

2. *To encode and implement for use by ONCOCIN the commonly used chemotherapy protocols from our oncology clinic.* In the upcoming year, we will:

   - Extend the OPAL protocol entry system

   - Continue entry of additional protocols, hopefully at the rate of one protocol/month (including testing)

3. *To continue testing of the workstation version of ONCOCIN.*

4. *To generalize the reasoning and interaction components of the ONCOCIN system for other applications.*

B. Justification and Requirements for Continued SUMEX Use

All the work we are doing (ONCOCIN plus continued use of the original MYCIN program) continues to be dependent on daily use of the SUMEX resource. Although much of the ONCOCIN work is shifting to Xerox workstations, the SUMEX 2060 and the 2020 continue to be key elements in our research plan. The programs all make assumptions regarding the computing environment in which they operate.

In addition, we have long appreciated the benefits of GUEST and network access to the programs we are developing. SUMEX greatly enhances our ability to obtain feedback from interested physicians and computer scientists around the country. Network access has also permitted high quality formal demonstrations of our work both from around the United States and from sites abroad (e.g., Finland, Japan, Sweden, Switzerland).

The main development of our project will continue to take place on LISP machines which we have purchased or which have been donated by the XEROX Corporation.

C. Requirements for Additional Computing Resources

The acquisition of the DEC 2020 by SUMEX was crucial to the growth of our research work. It has insured high quality demonstrations and has enabled us to develop a system (ONCOCIN) for real-world use in a clinical setting. As we have begun to develop systems that are potentially useful as stand-alone packages (i.e., an exportable ONCOCIN), the addition of personal workstations has provided particularly valuable new resources. We have made a commitment to the smaller Interlisp-D machines ("D-machines") produced by Xerox, and our work will increasingly transfer to them over the next several years. Our current funding supports our effort to implement ONCOCIN on workstations in the Stanford oncology clinic (and eventually to move the program to non-Stanford environments), but we will simultaneously continue to require access to Interlisp on upgraded workstations for extremely CPU-intensive tasks. Although our dependence on SUMEX for workstations has decreased due to a recent gift from XEROX, our requirements for network support of the machines has drastically increased. Individual machines do not provide sufficient space to store all of the software used in our project, nor to provide backup or long-term storage of work in progress. It is the networks, file storage devices, protocol converters, and other parts of the SUMEX network that hold our project together. In addition, with a research group of about 20 people, we are taking advantage of file sharing, electronic mail, and other information coordinating activities provided by the DEC 2060. We hope that with systems support and research by SUMEX staff, we will be able to gradually move away from a need for the central coordinating machine over the next five years.

The acquisition of the DEC 2060, coupled with our increasing use of workstations, has greatly helped with the problems in SUMEX response time that we had described in previous annual reports. We are extremely grateful for access both to the central machine and to the research workstations on which we are currently building the new ONCOCIN prototype. The D-machine's greater address space is permitting development of the large knowledge base that ONCOCIN requires. The graphics capability of the workstations has also enabled us to develop new methods for presenting material to naive users. In addition, the workstations have provided a reliable, constant "load-average" machine for running experiments with physicians and for development work. The development of ONCOCIN on the D-machine will demonstrate the feasibility of running intelligent consultation systems on small, affordable machines in physicians' offices and other remote sites.

D. Recommendations for Future Community and Resource Development

SUMEX is providing an excellent research environment and we are delighted with the help that SUMEX staff have provided implementing enhanced system features on the 2060 and on the workstations. We feel that we have a highly acceptable research environment in which to undertake our work. Workstation availability is becoming increasingly crucial to our research, and we have found over the past year that workstation access is at a premium. The SUMEX staff has been very helpful and understanding about our needs for workstation access, allowing us D-machine use wherever possible, and providing us with systems-level support when needed. We look forward to the arrival of additional advanced workstations and the development of a more distributed computing environment through SUMEX-AIM.

Responses to Questions Regarding Resource Future

"What do you think the role of the SUMEX-AIM resource should be for the period after 7/86, e.g., continue like it is, discontinue support of the central machine, act as a communications crossroads, develop software for user community workstations, etc.?"

We believe that the trend towards distributed computing that characterized the early

1980's will continue during the second half of the decade. Although we have begun this process by moving much of our research activity to LISP machines, the SUMEX DEC-20 continues to be a major source of support for all communication, collaboration, and administrative functions. It also continues to provide a quality LISP environment for rapid prototyping, student projects in the early stages before workstations are made available, and for demonstrating system features to people at a distance. These latter functions are still not well handled by distributed machines, and we believe that a logical role for the resource in the future is to develop software and communications techniques that will allow us to further decrease our dependence on the large central machine.

"Will you require continued access to the SUMEX-AIM 2060 and if so, for how long?"

As indicated above, our needs could still be met with a gradual phaseout of the 2060 over the next 3-5 years, provided that current services such as file handling and backup, mail, document preparation, and advanced network support are available from other machines (e.g., SAFE file server plus the Medical Computer Science file server). This implies maintenance of an ARPANET connection, connections to other campus machines, and facilities for linking together the heterogeneous collection of computing equipment upon which our research group depends. SUMEX would need to concentrate on providing software support for networks and systems software for workstations if it were to provide the same level of service we now experience while moving to a fully distributed environment.

"What would be the effect of imposing fees for using SUMEX resources (computing and communications) if NIH were to require this?"

Since all our research is NIH-supported, we see nothing but administrative headaches without benefits if there were to be a move to require fee-for-service billing for access to shared SUMEX resources. The net effect would simply be a transfer of funds from one arm of NIH to another (assuming that the agencies that currently fund our work could supplement our grants to cover SUMEX charges), and there would be a simultaneous restraining effect on the research environment. The current scheme permits experimentation and flexibility in use that would be severely inhibited if all access incurred an incremental charge.

"Do you have plans to move your work to another machine workstation and if so, when and to what kind of system?"

As mentioned above, and described in greater detail in our annual report, we are making a major effort to move much of our research activity to LISP machines (currently Xerox 1108's, 1186's and HP-9836's). Our familiarity with this technology, and our commitment to it, have resulted solely from the foresight of the SUMEX resource in anticipating the technology and providing for it at the time of their last renewal. However, for the reasons mentioned above, we continue to depend upon the central communication node for many aspects of our activities and could effectively adapt to its demise only if the phaseout were gradual and accompanied by improved support for a totally distributed computing environment.

IV.A.4. PROTEAN Project

PROTEAN Project

Oleg Jardetzky
Nuclear Magnetic Resonance Lab, School of Medicine
Stanford University

Bruce Buchanan, Ph.D.
Computer Science Department
Stanford University

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

The goals of this project are related both to biochemistry and artificial intelligence: (a) use existing AI methods to aid in the determination of the 3-dimensional structure of proteins in solution (not from x-ray crystallography proteins), and (b) use protein structure determination as a test problem for experiments with the AI problem solving structure known as the Blackboard Model. Empirical data from nuclear magnetic resonance (NMR) and other sources may provide enough constraints on structural descriptions to allow protein chemists to bypass the laborious methods of crystallizing a protein and using X-ray crystallography to determine its structure. This problem exhibits considerable complexity. Yet there is reason to believe that AI programs can be written that reason much as experts do to resolve these difficulties [8].

B. Medical Relevance

The molecular structure of proteins is essential for understanding many problems of medicine at the molecular level, such as the mechanisms of drug action. Using NMR data from proteins in solution will allow the study of proteins whose structure cannot be determined with other techniques, and will decrease the time needed for the determination.

C. Highlights of Progress

We have constructed a prototype of such a program, called PROTEAN, designed on the blackboard model [3] [4]. It is implemented in BB1 [5], a framework system for building blackboard systems that control their own problem-solving behavior [6] (see discussion of BB1 above). The reasoning component of PROTEAN directs the actions of the Geometry System (GS) [1], a set of programs that performs the computationally intensive task of positioning portions of a molecule with respect to each other in three dimensions. Currently we have implemented two versions of the GS: an InterLISP version used for quickly testing ideas and developing prototypes of geometric routines on a LISP workstation; and a high performance version written in C and running in the UNIX environment, providing efficient computations on a VAX 11/780. We have coupled the reasoning and geometry programs with an IRIS graphics terminal (shared with SUMEX) which displays the evolving protein structures at several levels of detail. The display in three dimensions is essential to understanding the behavior of the reasoning and geometry systems, and provides essential insights on the problem solving process.

PROTEAN embodies the following experimental techniques for coping with the complexities of constraint satisfaction:

1. The problem-solver partitions each problem into a network of loosely-coupled sub-problems. PROTEAN partitions the problem of positioning all of a protein's constituent structures within a global coordinate system into sub-problems of positioning individual pieces of structures and their immediate neighbors within local coordinate systems. It subsequently composes the most constrained partial solutions developed for these sub-problems in a complete solution for the entire protein. This partitioning and composition technique reduces the combinatorics of search. It also introduces additional constraints in the global characteristics of internally constrained partial solutions. For example, the conformations of partial protein solutions constrain their composability with other partial solutions. In addition, constraints on the overall dimensions of the protein from scattering experiments and indications of which atoms are on the surface of the molecule are used to further limit the possible structures.

2. The problem-solver attempts to solve sub-problems and coordinate solutions at multiple levels of abstraction, where lower levels of abstraction partition solution elements with finer granularity. For example, PROTEAN operates at three levels of abstraction. At the "Solid" level, it positions elements of the protein's secondary structure: alpha-helices, beta-sheets, and random coils. At the "Superatom" level, it positions elements of the protein's primary structure of amino acids: peptide units and side-chains. At the "Atom" level, it positions the protein's individual atoms. Partial solutions at higher levels of abstraction reduce the combinatorics of search at lower levels. Conversely, tightly constrained partial solutions at lower levels introduce new constraints on higher-level solutions.

3. The problem-solver forbears hypothesizing specific partial solutions for a sub-problem in favor of preserving the "family" of solutions consistent with all constraints applied thus far. For example, in positioning a helix within a partial solution, PROTEAN does not attempt to identify a unique spatial position for the helix. Instead, it identifies the entire spatial volume within which the helix might lie, given the constraints applied thus far. Preserving the family of legal solutions accommodates problems with incomplete constraints; the solution is only as constrained as the data are constraining. It also accommodates incompatible constraints by permitting disjunctive sub-families. For PROTEAN, disjunctive sub-volumes imply that the associated structure lies within any one of the sub-volumes or, if the structure is mobile, that it may move from one sub-volume to another.

4. The problem-solver applies constraints one at a time, successively restricting the family of solutions hypothesized for different sub-problems. PROTEAN successively applies constraints on the positions of protein structures, successively restricting the spatial volumes within which they may lie. Independent application of different constraints finesses the problem of integrating qualitatively different kinds of constraints by simply integrating their results. In addition, successive restriction of the family of solutions obviates guessing which specific solutions within a family are likely to be consistent with subsequently applied constraints and the otherwise inevitable back-tracking.

5. The problem-solver tolerates overlapping solutions for different sub-problems. For example, in identifying the volume within which structure-a might lie in partial solution 1, PROTEAN may include part of the volume identified for structure-b. Toleration of overlapping partial solutions is another accommodation of incomplete or incompatible constraints and

E. H. Shortliffe

potentially dynamic solutions. For PROTEAN, overlapping volumes for two protein structures indicate either: (a) that the two structures actually occupy disjoint sub-volumes that cannot be distinguished within the larger, overlapping volumes identified for them because the constraints are incomplete; or (b) that the two structures are mobile and alternately occupy the shared volume.

6. The problem-solver reasons explicitly about control of its own problem-solving actions: which sub-problems it will attack, which partial solutions it will expand, and which constraints it will apply. Control reasoning guides the problem-solver to perform actions that minimize computation, while maximizing progress toward a complete solution (see section 3.2.1). It also provides a foundation for the problem-solver's explanation of problem-solving activities and intermediate partial solutions (see section 3.2.2) and for its learning of new control heuristics (see section 5.5).

The current version of PROTEAN has five domain knowledge sources that demonstrate the reasoning techniques described above for the assembly of a protein. Each domain knowledge source directs a small portion of the construction of the molecule. These knowledge sources develop partial solutions that position alpha helices, beta strands, and random coils at the Solid level and refine the resulting state families using all available distance constraints.

PROTEAN now uses five control knowledge sources to guide the assembly of a solution to a protein structure. These knowledge sources determine which of the possible assembly actions is the best to perform at each stage of the problem solving.

Proposed work will introduce knowledge sources that apply the reasoning techniques for surface and volume constraints, as well as the ability to reason at Superatom and Atom levels. We also will investigate emergent constraints entailed in reliable partial solutions, composition of partial solutions into complete solutions, and intelligent control.

Multiple blackboards in PROTEAN allow several sets of knowledge to be used. A biochemical knowledge base stores information about proteins and secondary structures, amino acids, and atoms. The problem blackboard describes the protein to be solved and all experimental data observed for the molecule. The evolving solution of the protein structure is built on a third solution blackboard.

The PROTEAN system [2] has been used to construct a complete solution at the solid level of detail for the Lac-repressor headpiece, a protein with 51 amino acids consisting of 4 random coil sections and three alpha helices. In this work, the constraints were determined experimentally from NMR studies.

To demonstrate that our method is correct, we have applied PROTEAN to sperm whale myoglobin, a molecule whose crystal structure is known. By using distances between atoms in the crystal, distance constraints were applied to the eight helices in myoglobin to determine if PROTEAN would reproduce the crystal structure. The family of solutions obtained from PROTEAN includes the actual structure of the molecule. Work is proceeding to include the heme group of myoglobin as a component and use constraints to other portions of the molecule to further restrict the state families obtained.

D. Relevant Publications

1. Brinkley, J., Cornelius, C., Altman, R., Hayes-Roth, B. Lichtarge, O., Duncan, B., Buchanan, B.G., Jardetzky, O.: *Application of Constraint Satisfaction Techniques to the Determination of Protein Tertiary Structure.* Report KSL-86-28.

2. Buchanan, B.G., Altman, A., Brinkley, J., Cornelius, C., Duncan, B., Hayes-Roth, B., Hewett, M., Lichtarge, O., Jardetzky, O.: *The Heuristic Refinement Method for Deriving Solution Structures of Proteins*. Report KSL-85-41. October 1985. In *Proceedings of the National Academy of Sciences*.

3. Erman, L.D., Hayes-Roth, B., Lesser, V.R., Reddy, D.R.: *The HEARSAY-II Speech Understanding System: Integrating Knowledge to Resolve Uncertainty*. ACM Computing Surveys 12(2):213-254, June, 1980.

4. Hayes-Roth, B.: *The Blackboard Architecture: A General Framework for Problem Solving?* Report HPP-83-30, Department of Computer Science, Stanford University, 1983.

5. Hayes-Roth, B.: *BB1: An Environment for Building Blackboard Systems that Control, Explain, and Learn about their own Behavior*. Report HPP-84-16, Department of Computer Science, Stanford University, 1984.

6. Hayes-Roth, B.: *A Blackboard Architecture for Control*. Artificial Intelligence 26:251-321, 1985.

7. Hayes-Roth, B. and Hewett, M.: *Learning Control Heuristics in BB1*. Report HPP-85-2, Department of Computer Science, 1985.

8. Jardetzky, O.: *A Method for the Definition of the Solution Structure of Proteins from NMR and Other Physical Measurements: The LAC-Repressor Headpiece*. Proceedings of the International Conference on the Frontiers of Biochemistry and Molecular Biology, Alma Alta, June 17-24, 1984, October, 1984.

9. Hayes-Roth, B., Buchanan, B.G., Lichtarge, O., Hewitt, M., Altman, R., Brinkley, J., Cornelius, C., Duncan, B., and Jardetzky, O.: *PROTEAN: Deriving protein structure from constraints*. To appear in Proceedings of the AAAI, 1986.

E. Funding Support

Title: Interpretation of NMR Data from Proteins Using AI Methods

PI's: Oleg Jardetzky and Bruce G. Buchanan

Agency: National Science Foundation

Grant identification number: PCM 84-02348

Total Award Period and Amount: 11/1/84 - 10/31/86 $100,000
                 (includes direct and indirect costs)

Current award period and amount: 11/1/85 - 10/31/86 $ 50,000
                 (includes direct and indirect costs)

The following grants and contracts each provide partial funding for PROTEAN personnel.

Title: Research on Blackboard Problem-Solving Systems

PI: Edward A. Feigenbaum and Bruce G. Buchanan

Agency: Boeing Computer Services Corporation

Grant identification number: W-266875

Total award period and amount: 2/1/85 - 1/31/85, $225,000
        (direct and indirect)

Current award period and amount: 2/1/85 - 1/31/85, $225,000
        (direct and indirect)
        PROTEAN component is $11,400 (direct & indirect) or 5% of grant

Title: Knowledge-Based Systems Research

PI: Edward A. Feigenbaum

Agency: Defense Advanced Projects Research Agency

Grant identification number: N00039-86-0033

Total award period and amount: 10/1/85 - 9/30/88 $4,130,230 (in negotiation)
        (direct and indirect)

Current award period and amount: 10/1/85 - 9/30/86 $1,224,241
        (direct and indirect)
        PROTEAN component is $23175, or   1.9 % of grant total

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaborations

Several members of Prof. Jardetzky's research group are involved in this research.

B. Interactions with other SUMEX-AIM projects

Members of the PROTEAN project visits Robert Langridge's laboratory at the University of San Francisco last year, and informal discussions with him and his group have continued in this year.

C. Critique of Resource Management

The SUMEX staff has continued to be most cooperative in supporting PROTEAN research. The SUMEX computer facility is well maintained and managed for effective support of our work.

III. RESEARCH PLANS

A. Goals & Plans

Our long-range goal is to build an automatic interpretation system similar to CRYSALIS (which worked with x-ray crystallography data). In the shorter term, we are

building interactive programs that aid in the interpretation of NMR data on small proteins. The current version of PROTEAN has five domain and five control knowledge sources that demonstrate the reasoning techniques described above. These knowledge sources develop partial solutions that position multiple alpha helices, random coils, and beta structures at the Solid level and refine those helices using distance, surface, and volume constraints. The proposed research would expand PROTEAN to include knowledge sources that:

1. merge highly constrained partial solutions at the Solid level;

2. refine Solid level solutions in terms of the relative positions of constituent peptide units and side chains at the Superatom level;

3. further restrict the relative locations of peptide units and side chains relative to one another at the Superatom level;

4. propagate emergent constraints at the Superatom level back up to the Solid level to further restrict the relative positions of superordinate helices, beta sheets, and random coils;

5. refine Superatom level solutions at the Atom level;

6. further restrict the relative locations of atoms relative to one another;

7. propagate emergent constraints at the Atom level back up to the Superatom level to further restrict the relative positions of superordinate peptide units and side chains.

8. select instances of structures to be used as starting points for other kinds of refinement procedures, such as the solution of the Bloch equations, which define the NMR spectrum that can possibly arise from a given structure. These equations provide a very strong test of the correctness of our method, as well as providing an additional constraint on proposed structures.

9. position non-protein components with respect to partial solutions to proteins. Cofactors such as the heme group in myoglobin may be very constraining and lead to better structures.

10. develop efficient and effective control strategies for the solution of small, intermediate, and large molecules.

11. reason about mobility of structures when the data indicate that mobility is possible.

The research will also develop a set of control knowledge sources to guide PROTEAN's application of constraints to identify the family of legal protein conformations as efficiently as possible. We expect to improve the graphics interface to provide more functionality and options for viewing partial structures.

B. Justification for continued SUMEX use

We will continue to use SUMEX for developing parts of the program before integrating them with the whole system. We are using Interlisp to implement the Blackboard model and knowledge structures most flexibly and quickly.

C. Need for other computing resources

In this stage of development we need more computer cycles for geometric computations.

E. H. Shortliffe

We expect to upgrade the Silicon Graphics IRIS terminal to a workstation for more efficiency in the subprograms doing computational geometry and for more effect graphical display of the results of PROTEAN.

IV.A.5. RADIX Project

The RADIX Project: Deriving Medical Knowledge from
Time-Oriented Clinical Databases

Robert L. Blum, M.D., Ph.D.
Department of Computer Science
Stanford University

Gio C. M. Wiederhold, Ph.D.
Departments of Computer Science and Medicine
Stanford University

I.   SUMMARY OF RESEARCH PROGRAM

A. Technical Goals - Introduction

Medical and Computer Science Goals -- The long-range objectives of our project, called RADIX (formerly RX), are 1) to increase the validity of medical knowledge derived from large time-oriented databases containing routine, non-randomized clinical data, 2) to provide knowledgeable assistance to a research investigator in studying medical hypotheses on large databases, 3) to fully automate the process of hypothesis generation and exploratory confirmation. For system development we have used a subset of the ARAMIS database.

Computerized clinical databases and automated medical records systems have been under development throughout the world for at least a decade. Among the earliest of these endeavors was the ARAMIS Project, (American Rheumatism Association Medical Information System) under development since 1969 in the Stanford Department of Medicine. ARAMIS contains records of over 17,000 patients with a variety of rheumatologic diagnoses. Over 62,000 patient visits have been recorded, accounting for 50,000 patient-years of observation. The ARAMIS Project has now been generalized to include databases for many chronic diseases other than arthritis.

The fundamental objective of the ARAMIS Project and many other clinical database projects is to use the data that have been gathered by clinical observation in order to study the evolution and medical management of chronic diseases. Unfortunately, the process of reliably deriving knowledge has proven to be exceedingly difficult. Numerous problems arise stemming from the complexity of disease, therapy, and outcome definitions, from the complexity of causal relationships, from errors introduced by bias, and from frequently missing and outlying data. A major objective of the RADIX Project is to explore the utility of symbolic computational methods and knowledge-based techniques at solving some of these problems.

The RADIX computer program is designed to examine a time-oriented clinical database such as ARAMIS and to produce a set of (possibly) causal relationships. The algorithm exploits three properties of causal relationships: time precedence, correlation, and nonspuriousness. First, a Discovery Module uses lagged, nonparametric correlations to generate an ordered list of tentative relationships. Second, a Study Module uses a knowledge base (KB) of medicine and statistics to try to establish nonspuriousness by controlling for known confounders.

The principal innovations of RADIX are the Study Module and the KB. The Study

Module takes a causal hypothesis obtained from the Discovery Module and produces a comprehensive study design, using knowledge from the KB. The study design is then executed by an on-line statistical package, and the results are automatically incorporated into the KB. Each new causal relationship is incorporated as a machine-readable record specifying its intensity, distribution across patients, functional form, clinical setting, validity, and evidence. In determining the confounders of a new hypothesis the Study Module uses previously "learned" causal relationships.

In creating a study design the Study Module follows accepted principles of epidemiological research. It determines study feasibility and study design: cross-sectional versus longitudinal. It uses the KB to determine the confounders of a given hypothesis, and it selects methods for controlling their influence: elimination of patient records, elimination of confounding time intervals, or statistical control. The Study Module then determines an appropriate statistical method, using knowledge stored as production rules. Most studies have used a longitudinal design involving a multiple regression model applied to individual patient records. Results across patients are combined using weights based on the precision of the estimated regression coefficient for each patient.

More recently, we have undertaken two new components to the RADIX program: a module for automated summarization of patient records, and a knowledge-based discovery module. The goal of the summarization program is to automatically create patient summaries of arbitrary and appropriate complexity as an aid for tasks such as clinical decision making, real-time patient monitoring, surveillance of quality of care, and eventually automated discovery. This program builds on our experience with labelling patient records in RX, and is a natural extension of our work on the interface between AI and medical databases. The goal of the knowledge-based discovery module is to overcome some of the limitations of the original, statistics-based, RX discovery module. In creating disease hypotheses, researchers make extensive use of notions of causation, mechanism of action, tempo, and quantitative sufficiency, as well as detailed knowledge of pathophysiology. We are seeking to automate this process of hypothesis formation by replicating selected discoveries in rheumatology using data from the ARAMIS database.

B. Medical Relevance and Collaboration

As a test bed for system development, our focus of attention has been on the records of patients with systemic lupus erythematosus (SLE) contained in the Stanford portion of the ARAMIS Data Bank. SLE is a chronic rheumatologic disease with a broad spectrum of manifestations. Occasionally the disease can cause profound renal failure and lead to an early death. With many perplexing diagnostic and therapeutic dilemmas, it is a disease of considerable medical interest.

In the future we anticipate possible collaborations with other project users of the TOD System such as the National Stroke Data Bank, the Northern California Oncology Group, and the Stanford Divisions of Oncology and of Radiation Therapy.

We believe that this research project is broadly applicable to the entire gamut of chronic diseases that constitute the bulk of morbidity and mortality in the United States. Consider five major diagnostic categories responsible for approximately two thirds of the two million deaths per year in the United States: myocardial infarction, stroke, cancer, hypertension, and diabetes. Therapy for each of these diagnoses is fraught with controversy concerning the balance of benefits versus costs.

1. Myocardial Infarction: Indications for and efficacy of coronary artery bypass graft vs. medical management alone. Indications for long-term antiarrhythmics ... long-term anticoagulants. Benefits of cholesterol-lowering diets, exercise, and so forth.

2. Stroke: Efficacy of long-term anti-platelet agents, long-term anticoagulation. Indications for revascularization.

3. Cancer: Relative efficacy of radiation therapy, chemotherapy, surgical excision - singly or in combination. Optimal frequency of screening procedures. Prophylactic therapy.

4. Hypertension: Indications for therapy. Efficacy versus adverse effects of chronic antihypertensive drugs. Role of various diagnostic tests such as renal arteriography in work-up.

5. Diabetes: Influence of insulin administration on microvascular complications. Role of oral hypoglycemics.

Despite the expenditure of billions of dollars over recent years for randomized controlled trials (RCT's) designed to answer these and other questions, answers have been slow in coming. RCT's are expensive in terms of funds and personnel. The therapeutic questions in clinical medicine are too numerous for each to be addressed by its own series of RCT's.

On the other hand, the data regularly gathered in patient records in the course of the normal performance of health care delivery are a rich and largely underutilized resource. The ease of accessibility and manipulation of these data afforded by computerized clinical databases holds out the possibility of a major new resource for acquiring knowledge on the evolution and therapy of chronic diseases.

The goal of the research that we are pursuing on SUMEX is to increase the reliability of knowledge derived from clinical data banks with the hope of providing a new tool for augmenting knowledge of diseases and therapies as a supplement to knowledge derived from formal prospective clinical trials. Furthermore, the incorporation of knowledge from both clinical data banks and other sources into a uniform knowledge base should increase the ease of access by individual clinicians to this knowledge and thereby facilitate both the practice of medicine as well as the investigation of human disease processes.

The medical relevance of the automated summarization program is readily apparent. A practicing physician or medical researcher, faced with a patient chart, often with dozens of visits and scores of attributes, rarely has time to read the entire chart. He (or she) would like a succinct summary of the important events in that patient's record to assist his decision making. The objective of our current work is to implement an artificial-intelligence based computer program that produces such summaries.

C. Highlights of Research Progress

C.1 April 1985 to April 1986

Our primary accomplishments in this period have been the following:

1) Design and implementation of a prototype automated summarization program.

2) Design of an intelligent medical hypothesis generator for the discovery module, partially implemented.

3) Identification and comparison of novel Exploratory Data Analysis (EDA) methods for use in automated discovery.

4) Installation of the KEE Knowledge Engineering Environment on the 1108's and training of project members in KEE.

5) Publication of papers on automated discovery and automated summarization, and presentation of results at medical conferences.

6) Training Post-Doctoral researchers, participants in RADIX, in methods of medical artificial intelligence research.

C.1.1 Design and implementation of a prototype automated summarization program

We have designed and implemented a prototype for the automated summarization program using KEE. The prototype labels and summarizes a small number of major events in time-oriented medical records of systemic lupus erythematosus patients. The user interface provides an interactive, graphic representation of the record with active regions selectable by the user to display data supporting conclusions or to magnify a region showing greater detail of the patient record. The program uses a hypothetico-deductive algorithm. Disease states are evoked based on attributes with abnormal values, and the evoked disease frames are confirmed by matching their templates against the patient record using a temporal querying syntax. This produces likelihood ratios that are used for Bayesian updating of the evoked disease states. A knowledge base of definitions of medical objects and their causal relations, implemented in KEE, underlies the program. This work is described in Downs, 1986, noted in the publications section.

C.1.2 Design of an intelligent medical hypothesis generator for the Discovery Module

We have designed an intelligent medical hypothesis generator for the Discovery Module, which we are currently implementing. The design will evolve as we gain experience with the system. In contrast with the original RX Discovery Module, which relied on statistical methods, the new program relies on artificial intelligence methods. In this work we are interested in elaborating the cognitive mechanisms whereby disease hypotheses are formulated in terms of clinical events and the known inter-relationships among agents of disease and organ function. Clearly, in creating disease hypotheses researchers make extensive use of plausible notions of causation, mechanism of action, tempo, and quantitative sufficiency, and so on.

Our overall methodology involves several stages: 1) analysis of important, previous discoveries from the biomedical literature relevant to our patient database, 2) formulation of a theory of how these discoveries could have been made, 3) development of knowledge representation and reasoning mechanisms adequate to embody the theory in software, 4) computer simulation of the selected discoveries to evaluate and refine the implementation, 5) generalization of the program to other selected discoveries, and 6) operation of the program in a self-guided mode to seek previously unknown findings. This work is described in Walker, 1986, noted in the publications.

C.1.3 Identification and comparison of novel Exploratory Data Analysis (EDA) methods for use in automated discovery

The field of Exploratory Data Analysis (EDA) is one of the roots of our work in automated discovery. Our work on statistical methods for hypothesis generation has sought to extend our earlier work on the RX Discovery Module, which was based on nonparametric correlations on all variables at several time lags. There are many cases in which data will show no correlation, but still have an interesting structure. For example, if the data fall into clusters, are non-linear, or follow a regular pattern they may suggest an important relationship but have zero correlation. We have identified and compared a number of novel EDA methods for use in these situations, described in Walker, 1986, noted in the publications section.

C.1.4 Installation of the KEE Knowledge Engineering Environment on the 1108's and training of project members in KEE

We have been fortunate in obtaining use of the KEE Knowledge Engineering

Environment, a very valuable tool for aiding development of artificial intelligence software. KEE, from Intellicorp, costs approximately $35,000, but was supplied to us without charge. KEE is installed on our machines, and project members are now using KEE for program development.

C.1.5 Publication of papers on automated discovery and automated summarization, and presentation of results at medical conferences

In addition to the publications noted above, we have submitted and/or had accepted additional papers, noted in the section on publications, and presented results at numerous medical conferences.

C.1.6 Training Post-Doctoral researchers, participants in RADIX, in methods of medical artificial intelligence research

We have been training three post-doctoral researchers on the project during the current reporting year; Steven Downs, M.D., Isabelle de Zegher-Geets, M.D., and Donald Rucker, M.D.. Steven Downs will complete a thesis as part of Stanford's Medical Information Sciences program this June; Rucker and de Zegher-Geets will undertake theses in the coming year.

C.2 Research in Progress

Our current research carries forward the work in automated summarization and automated discovery described above. Specifically, we are 1) implementing the intelligent discovery module, and evaluating and modifying its design as we get initial results, and 2) substantially expanding the prototype automated summarization module to be able to deal with a full patient record. We continue to work on problems involved in the representation of medical knowledge, as part of developing the programs for summarization and discovery. These programs act both as test beds for the extant knowledge representation techniques, and forcing functions for the development of new techniques.

D. Publications

1. Blum, R.L. and Walker, M. G.: *Lisp as an Environment for Software Design* Invited paper for Tenth Annual Symposium on Computers in Medical Care (SCAMC), Washington, D.C., October 1986.

2. Blum, R.L.: *Computer-Assisted Design of Studies using Routine Clinical Data: Analyzing the Association of Prednisone and Cholesterol.* (Accepted for publication in the Annals of Internal Medicine.)

3. Blum, Robert L. and Gio C.M. Wiederhold: *Studying Hypotheses on a Time-Oriented Clinical Database: An Overview of the RX Project.* In J.A.Reggia and S.Thurim: 'Computer Assisted Medical Decision-Making'; Springer Verlag, 1985, pp.245--253.

4. Blum, R.L.: *Two Stage Regression: Application to a Time-Oriented Clinical Database.* Knowledge Systems Laboratory Technical Report. 1985.

5. Blum, R.L.: *Modeling and encoding clinical causal relationships.* Proceedings of SCAMC, Baltimore, MD, October, 1983.

6. Blum, R.L.: *Representation of empirically derived causal relationships.* IJCAI, Karlsruhe, West Germany, August, 1983 .

7. Blum, R.L.: *Machine representation of clinical causal relationships.* MEDINFO 83, Amsterdam, August, 1983.

8. Blum, R.L.: *Clinical decision making aboard the Starship Enterprise.* Chairman's paper, Session on Artificial Intelligence and Clinical Decision Making, AAMSI, San Francisco, May, 1983.

9. Blum, R.L. and Wiederhold, G.: *Studying hypotheses on a time-oriented database: An overview of the RX project.* Proc. Sixth SCAMC, IEEE, Washington D.C., October, 1982.

10. Blum, R.L.: *Induction of causal relationships from a time-oriented clinical database: An overview of the RX project.* Proc. AAAI, Pittsburgh, August, 1982.

11. Blum, R.L.: *Automated induction of causal relationships from a time-oriented clinical database: The RX project.* Proc. AMIA San Francisco, 1982.

12. Blum, R.L.: *Discovery and Representation of Causal Relationships from a Large Time-oriented Clinical Database: The RX Project.* In D.A.B. Lindberg and P.L. Reichertz (Eds.), LECTURE NOTES IN MEDICAL INFORMATICS, Springer-Verlag, 1982.

13. Blum, R.L.: *Discovery, confirmation, and incorporation of causal relationships from a large time-oriented clinical database: The RX project.* Computers and Biomed. Res. 15(2):164-187, April, 1982.

14. Blum, R.L.: *Discovery and representation of causal relationships from a large time-oriented clinical database: The RX project* (Ph.D. thesis). Computer Science and Biostatistics, Stanford University, 1982.

15. Blum, R.L.: *Displaying clinical data from a time-oriented database.* Computers in Biol. and Med. 11(4):197-210, 1981.

16. Blum, R.L.: *Automating the study of clinical hypotheses on a time-oriented database: The RX project.* Proc. MEDINFO 80, Tokyo, October, 1980, pp. 456-460. (Also STAN-CS-79-816)

17. Blum, R.L. and Wiederhold, G.: *Inferring knowledge from clinical data banks utilizing techniques from artificial intelligence.* Proc. Second SCAMC, IEEE, Washington, D.C., November, 1978.

18. Blum, R.L.: *The RX project: A medical consultation system integrating clinical data banking and artificial intelligence methodologies,* Stanford University Ph.D. thesis proposal, August, 1978.

19. Downs, S., Walker, M.G. and Blum, R.L.: *Automated Summaries of Patient Medical Records.* Accepted for Medinfo '86.

20. Kuhn, I., Wiederhold, G., Rodnick, J.E., Ramsey-Klee, D.M., Benett, S., Beck, D.D.: *Automated Ambulatory Medical Record Systems in the U.S.,* to be published by Springer-Verlag, 1983, in *Information Systems for Patient Care,* B. Blum (ed.), Section III, Chapter 14.

21. Walker, M.G.: *How Feasible is Automated Discovery?* Knowledge Systems Laboratory Technical Report KSL 86-35. Submitted to IEEE Expert.

22. Walker, M.G., and Blum, R.L.: *Towards Automated Discovery from Clinical Databases: The RADIX Project* Accepted for Medinfo '86.

23. Walker, M.G., Blum, R.L., and Fagan, L.M.: *Minimycin: A Miniature Rule-Based System.* M.D.Computing, Vol. 2, No. 4., 1985.

24. Walker, M.G., and Blum, R.L.: *An Introduction to LISP.* M.D.Computing, Vol. 2, No. 1., 1985.

25. Wiederhold, Gio, Robert L. Blum, and Michael Walker: *An Integration of Knowledge and Data Representation.* Proc. of Islamorada Workshop, Feb.1985, Computer Corporation of America, Cambridge MA; Chapter 23 of 'On Knowledge Base Management Systems: Integrating Artificial Intelligence and Database Technologies' (Brodie, Mylopoulos, and Schmidt, eds.), Springer Verlag, June 1986.

26. Wiederhold, Gio: *Knowledge versus Data.* Chapter 9 of 'On Knowledge Base Management Systems: Integrating Artificial Intelligence and Database Technologies' (Brodie, Mylopoulos, and Schmidt, eds.) Springer Verlag, Feb.1986.

27. Missikoff, Michele and Gio Wiederhold: *Towards a Unified Approach for Expert and Database Systems.* In 'Expert Database Systems', Larry Kerschberg (editor), Benjamin/Cummings, 1986, pages 383-399; also in Proceedings of First Workshop on Expert Database Systems, Kiawah Island, South Carolina, Oct.1984, vol.1, pp.186-206.

28. Wiederhold, Gio and Paul D. Clayton: *Processing Biological Data in Real Time.* M.D. Computing, Springer Verlag, Vol.2 No.6, November 1985, pages 16-25.

29. Wiederhold, Gio: *Knowledge Bases.* Future Generations Computer Systems, North-Holland, vol.1 no.4; April 1985, pp.223--235.

30. Wiederhold, Gio: *Disease Registers: Use of Databases to Generate New Medical Knowledge.* In K.Abt, W.Giere and B.Leiber: 'Krankendaten, Krankheitsregister, Datenschutz', vol.58, Medizinische Informatik und Statistik, Springer Verlag, 1985, pp.39-55.

31. Wiederhold, G.: *Networking of Data Information,* National Cancer Institute Workshop on the Role of Computers in Cancer Clinical Trials, National Institutes of Health, June 1983, pp.113-119.

32. Wiederhold, G.: *Database Design* (in the Computer Science Series) McGraw-Hill Book Company, New York, NY, May 1977, 678 pp. Second edition, Jan. 1983, 768 pp.

33. Wiederhold, G.: In D.A.B. Lindberg and P.L. Reichertz (Eds.), *Databases for Health Care,* Lecture Notes in Medical Informatics, Springer-Verlag, 1981.

34. Wiederhold, G.: *Database technology in health care.* J. Medical Systems 5(3):175-196, 1981.

E. Funding Support Status

E. H. Shortliffe

Term: March 15, 1984 through March 14, 1986

2) Deriving Knowledge from Clinical Databases
   Gio C. M. Wiederhold, Ph.D.: Principal Investigator
   National Library of Medicine: LM-04334
   Total award: $291,192 (direct)
   Term: May 1, 1984 through November 30, 1986

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Collaborations

Once the RADIX program is developed, we would anticipate collaboration with some of the ARAMIS project sites in the further development of a knowledge base pertaining to the chronic arthritides. The ARAMIS Project at the Stanford Center for Information Technology is used by a number of institutions around the country via commercial leased lines to store and process their data. These institutions include the University of California School of Medicine, San Francisco and Los Angeles; The Phoenix Arthritis Center, Phoenix; The University of Cincinnati School of Medicine; The University of Pittsburgh School of Medicine; Kansas University; and The University of Saskatchewan. All of the rheumatologists at these sites have closely collaborated with the development of ARAMIS, and their interest in and use of the RADIX project is anticipated. We hasten to mention that we do not expect SUMEX to support the active use of RADIX as an on-going service to this extensive network of arthritis centers, but we would like to be able to allow the national centers to participate in the development of the arthritis knowledge base and to test that knowledge base on their own clinical data banks.

B. Interactions with Other SUMEX-AIM Projects

During the current reporting year we have had frequent interaction with members of other Sumex projects; for example, to discuss theoretical issues in discovery and automated summarization, practical programming issues, and to assist training of Medical Computer Science Students in the use of KEE, Lisp workstations, and so on. The Sumex community is an invaluable resource for providing such interaction.

C. Critique of Resource Management

The DEC System 20 continues to provide acceptable performance, but it is frequently heavily loaded at peak hours.

The SUMEX resource management continues to be accessible and quite helpful.

III. RESEARCH PLANS

A. Project Goals and Plans

The overall goal of the RADIX Project is to develop a computerized medical information system capable of accurately extracting medical knowledge pertaining to the therapy and evolution of chronic diseases from a database consisting of a collection of stored patient records.

SHORT-TERM GOALS --

Our short term goals focus on the two activities described earlier: implementation and

further development of the intelligent discovery module, and substantial expansion of the automated summarization program to deal with an entire rheumatology patient record.

LONG-RANGE GOALS --

The long-range goals of the RADIX Project are 1) automatic discovery of knowledge in a large time-oriented database, and provision of assistance to a clinician who is interested in testing a specific hypothesis, and 2) development of techniques for automated summarization of patient records. We hope to make these programs sufficiently robust that they will work over a broad range of hypotheses and over a broad spectrum of patient records.

B. *Justification and Requirements for Continued Use of SUMEX*

Computerized clinical data banks possess great potential as tools for assessing the efficacy of new diagnostic and therapeutic modalities, for monitoring the quality of health care delivery, and for support of basic medical research. Because of this potential, many clinical data banks have recently been developed throughout the United States. However, once the initial problems of data acquisition, storage, and retrieval have been dealt with, there remains a set of complex problems inherent in the task of accurately inferring medical knowledge from a collection of observations in patient records. These problems concern the complexity of disease and outcome definitions, the complexity of time relationships, potential biases in compared subsets, and missing and outlying data. The major problem of medical data banking is in the reliable inference of medical knowledge from primary observational data.

We see in the RADIX Project a method of solution to this problem through the utilization of knowledge engineering techniques from artificial intelligence. The RADIX Project, in providing this solution, will provide an important conceptual and technological link to a large community of medical research groups involved in the treatment and study of the chronic arthritides throughout the United States and Canada, who are presently using the ARAMIS Data Bank through the CIT facility via TELENET.

Beyond the arthritis centers which we have mentioned in this report, the TOD (Time-Oriented Data Base) User Group involves a broad range of university and community medical institutions involved in the treatment of cancer, stroke, cardiovascular disease, nephrologic disease, and others. Through the RADIX Project, the opportunity will be provided to foster national collaborations with these research groups and to provide a major arena in which to demonstrate the utility of artificial intelligence to clinical medicine.

C. *Recommendations for Resource Development*

The on-going acquisition of personal work-station Lisp processors is a very positive step, as these provide an excellent environment for program development, and can serve as a vehicle for providing programs to collaborators at other sites. Continued acquisitions are very desirable.

We also would hope that the central SUMEX facility, the DEC 2060, would continue to be supported. We continue to make constant use of this machine for text-editing, document preparation, file and database handling, communications, and program demos.

Responses to Questions Regarding Resource Future

Q: What do you think the role of the SUMEX-AIM resource should be for the period after 7/86, e.g., continue like it is,

E. H. Shortliffe

discontinue support of the central machine, act as a
communications crossroads, develop software for user
community workstations, etc.

A: In our opinion, the SUMEX 2060 should continue to be
      supported. The machine continues to be of value to us for
      text-editing (TV edit and EMACS) and for document preparation
      (SCRIBE) and for communications and mail. We also depend on
      it as a central, reliable facility for program demos, for
      manipulating large databases, and maintaining central program
      files. It would be a real loss if it was discontinued.

      Software for community work stations. Yes. Making good utility
      programs available to all users sounds like a good idea.

Q: Will you require continued access to the SUMEX-AIM 2060 and
      if so, for how long?

A: Yes. For the foreseeable future and for the above reasons.

Q: What would be the effect of imposing fees for using SUMEX
      resources (computing and communications) if NIH were to
      require this?

A: We would pay them. The 2060 is worth it to us. Of course,
      if the fees were high, we would consider alternatives.

Q: Do you have plans to move your work to another machine
      workstation and if so, when and to what kind of system?

A: We are currently using two of the SUMEX Xerox 1108's for
      the development of our project. We will stay with these
      for the foreseeable future.

IV.B. National AIM Projects

The following group of projects is formally approved for access to the AIM aliquot of the SUMEX-AIM resource. Their access is based on review by the AIM Advisory Group and approval by the AIM Executive Committee.

In addition to the progress reports presented here, abstracts for each project and its individual users are submitted on a separate Scientific Subproject Form.

E. H. Shortliffe

IV.B.1. INTERNIST-I Project

CADUCEUS Project (INTERNIST-I)

This project is unfunded at the present time.

J. D. Myers, M.D.
University Professor Emeritus (Medicine)
University of Pittsburgh
1291 Scaife Hall
Pittsburgh, Pa., 15261

I. SUMMARY OF RESEARCH PROGRAM

A. Project rationale

The principal objective of this project is the development of a high-level computer diagnostic program in the broad field of internal medicine as an aid in the solution of complex and complicated diagnostic problems. To be effective, the program must be capable of multiple diagnoses (related or independent) in a given patient.

A major achievement of this research undertaking has been the design of a program called INTERNIST-1, along with an extensive medical knowledge base. This program has been used over the past decade to analyze many hundreds of difficult diagnostic problems in the field of internal medicine. These problem cases have included cases published in medical journals (particularly Case Records of the Massachusetts General Hospital, in the New England Journal of Medicine), CPCs, and unusual problems of patients in our Medical Center. In most instances, but by no means all, INTERNIST-I has performed at the level of the skilled internist, but the experience has highlighted several areas for improvement.

B. Medical Relevance and Collaboration

The program inherently has direct and substantial medical relevance.

The development of the QUICK MEDICAL REFERENCE (QMR) under the leader ship of Dr. Randolph A. Miller has allowed us to distribute the INTERNIST-I knowledge base in a modified format to over 20 other academic medical institutions. The knowledge base can thereby be used as an "electronic textbook" in medical education at all levels -- by medical students, residents and fellows, and faculty and staff physicians. This distribution is continuing to expand.

The INTERNIST-I program has been used in recent years to develop patient management problems for the American College of Physician's Medical Knowledge Self-assessment Program, and to develop patient management problems and test cases for the Part III Examination and the developing computerized testing program of the National Board of Medical Examiners.

C. Highlights of Research Progress

C.1 Accomplishments this past year

The group of us (Myers, Miller and Masarie) together with assigned residents in internal

medicine are continuing to expand the knowledge base and to incorporate the diagnostic consultative program into QMR. The computer program for the interrogative part of the diagnostic program is the main remaining task. An editor for the QMR knowledge base, as modified from the INTERNIST-I knowledge base, is nearing completion. The entire QMR program can be accommodated in, maintained (particularly edited) and operated on individual IBM PC-AT computers.

In the near future our group will be ready to incorporate into the QMR diagnostic consultant program the modifications and embellishments of the INTERNIST-I knowledge base, e.g. the use of "facets" of diseases or syndromes. This addition and modification is expected to improve the performance of the diagnostic consultant program.

The medical knowledge base has continued to grow both in the incorporation of new diseases and the modification of diseases already profiled so as to include recent advances in medical knowledge. Several dozen new diseases have been profiled during the past year.

C.2 Research in progress

There are four major components to the continuation of this research project:

1. The enlargement, continued updating, refinement and testing of the extensive medical knowledge base required for the operation of INTERNIST-I and the QMR modification.

2. Institution of field trials of QMR on the clinical services in internal medicine at the Health Center of the University of Pittsburgh.

3. Expansion of the clinical field trials to other university health centers which have expressed interest in working with the system.

4. Adaptation of the diagnostic program and data base of INTERNIST-I and the QMR modification to subserve educational purposes and the evaluation of clinical performance and competence.

Current activity is devoted mainly to the first of these, namely, the continued development of the medical knowledge base, and the implementation of the improved diagnostic consulting program.

D. List of relevant publications

1. Myers, J.D.: Educating future physicians: Something old, Something new. Ohio State Univ. Proceedings of Symposium, Medical Education in the 21st Century. 1985.

2. Myers, J.D.: The process of clinical diagnosis and its adaptation to the computer IN The Logic of Discovery and Diagnosis in Medicine, University of Pittsburgh Series in the Philosophy and History of Science, edited by Kenneth F. Schaffner, Univ. of California Press, pp. 155-180, 1985.

3. Masarie, Jr. F.E., Miller, R.A., First, M.B., Myers, J.D.: An Electronic Textbook of Medicine. Proceedings of Ninth Annual Symposium on Computer Applications in Medical Care. Baltimore, Maryland, November 1985.

4. Masarie, Jr. F.E., Myers, J.D., Miller, R.A.: INTERNIST-I PROPERTIES: Representing Common Sense on Good Medical Practice in a Computerized

Medical Knowledge Base.   Computers and Biomedical Research.   18:
458-479, October 1985.

5. Myers, J.D., Chairman.   Medical Education in the Information Age.
Proceedings of the Symposium on Medical Informatics.   Association of
American Medical Colleges, 1986.

E. Funding support

1. Clinical Decision Systems Research Resource
   Harry E. Pople, Jr., Ph.D.
   Professor of Business
   Jack D. Myers, M.D.
   University Professor Emeritus (Medicine)
   University of Pittsburgh
   Division of Research Resources
   National Institutes of Health

   5 R24 RR01101-08
   07/01/80 - 03/31/86 - $1,658,347
   07/01/84 - 09/30/85 - $354,211
   09/30/85 - 03/31/86 - $50,690

2. CADUCEUS: A Computer-Based Diagnostic Consultant
   Harry E. Pople, Jr., Ph.D.
   Professor of Business
   Jack D. Myers, M.D.
   University Professor Emeritus (Medicine)
   University of Pittsburgh
   National Library of Medicine
   National Institutes of Health

   5 R01 LM03710-05
   07/01/80 - 03/31/86 - $853,200
   07/01/84 - 09/30/85 - $210,091
   09/30/85 - 03/31/86 - $35,316

3. Diagnostic-Internist: A Computerized Medical Consultant
   Randolph A. Miller, M.D.
   Associate Professor of Medicine
   University of Pittsburgh Department of Medicine
   National Library of Medicine - Development Award Research
   Career
   National Institutes of Health

   1 K04 LM00084-01
   09/30/85 - 09/29/90 - amounts to be determined annually
   09/30/85 - 09/29/86 - $55,296

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A,B. Medical Collaborations and Program Dissemination Via SUMEX

INTERNIST-I and QMR remains in a stage of research and particularly development.
As noted above, we are continuing to develop better computer programs to operate the
diagnostic system, and the knowledge base cannot be used very effectively for
collaborative purposes until it has reached a critical stage of completion. These factors
have stifled collaboration via SUMEX up to this point and will continue to do so for
the next year or two. In the meanwhile, through the SUMEX community there

continues to be an exchange of information and states of progress. Such interactions particularly take place at the annual AIM Workshop.

C. *Critique of Resource Management*

SUMEX has been an excellent resource for the development of INTERNIST-I. Our large program is handled efficiently, effectively and accurately. The staff at SUMEX have been uniformly supportive, cooperative, and innovative in connection with our project's needs.

III. RESEARCH PLANS

A. *Project Goals and Plans*

Continued effort to complete the medical knowledge base in internal medicine will be pursued including the incorporation of newly described diseases and new or altered medical information on "old" diseases. The latter two activities have proven to be more formidable than originally conceived. Profiles of added diseases plus other information is first incorporated into the medical knowledge base at SUMEX before being transferred into our newer information structures for QMR. This sequence retains the operative capability of INTERNIST-I as a computerized "textbook of medicine" for educational purposes.

B. *Justification and Requirements for Continued SUMEX Use*

Our use of SUMEX will obviously decline with the adaptation of our programs to the IBM PC-AT. Nevertheless, the excellent facilities of SUMEX are expected to be used for certain developmental work. It is intended for the present to keep INTERNIST-1 at SUMEX for comparative use as QMR is developed here.

Our best prediction is that our project will require continued access to the 2060 for the next year or two and we consider such access essential to the future development of our knowledge base. After that time, our work can probably be accomplished on our personal work stations.

C. *Needs and Plans for Other Computing Resources Beyond SUMEX-AIM*

Our predictable needs in this area will be met by our newly acquired personal work stations.

E. H. Shortliffe

IV.B.2. CLIPR - Hierarchical Models of Human Cognition

Hierarchical Models of Human Cognition (CLIPR Project)

Walter Kintsch and Peter G. Polson
University of Colorado
Boulder, Colorado

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

The two CLIPR projects have made progress during the last year. The prose comprehension project has completed one major project, and is designing a prose comprehension model that reflects state-of-the-art knowledge from psychology (van Dijk & Kintsch, 1983) and artificial intelligence. During the last four years, Polson, in collaboration with Dr. David Kieras of the University of Michigan, has continued work on a project studying the psychological factors underlying device complexity and the difficulties that nontechnically trained individuals have in learning to use devices like word processors. They have developed formal representations of a user's knowledge of how to operate a device and of the user-device interface (Kieras & Polson, 1985) and have completed several experiments evaluating their theory (Polson & Kieras, 1984, 1985; Polson, Muncher, and Engelbeck, 1986).

B. Technical Goals

The CLIPR project consists of two subprojects. The first, the text comprehension project, is headed by Walter Kintsch and is a continuation of work on understanding of connected discourse that has been underway in Kintsch's laboratory for several years. The second, the device complexity project, is headed by Peter Polson in collaboration with David Kieras of the University of Michigan. They are studying the learning and problem solving processes involved in the utilization of devices like word processors or complex computer controlled medical instruments (Kieras & Polson, 1985).

The goal of the prose comprehension project is to develop a computer system capable of the meaningful processing of prose. This work has been generally guided by the prose comprehension model discussed by van Dijk & Kintsch (1983), although our programming efforts have identified necessary clarifications and modifications in that model (Kintsch & Greeno, 1985; Fletcher, 1985; Walker & Kintsch, 1985; Young, 1985). In general, this research has emphasized the importance of knowledge and knowledge-based processes in comprehension. We hope to be able to merge the substantial artificial intelligence research on these systems with psychological interpretations of prose comprehension, resulting in a computational model that is also psychologically respectable.

The goal of the device complexity project is to develop explicit models of the user-device interaction. They model the device as a nested automata and the user as a production system. These models make explicit kinds of knowledge that are required to operate different kinds of devices and the processing loads imposed by different implementations of a device.

C. Medical Relevance and Collaboration

The text comprehension project impacts indirectly on medicine, as the medical

profession is no stranger to the problems of the information glut. By adding to the research on how computer systems might understand and summarize texts, and determining ways by which the readability of texts can be improved, medicine can only be helped by research on how people understand prose. Development of a more thorough understanding of the various processes responsible for different types of learning problems in children and the corresponding development of a successful remediation strategy would also be facilitated by an explicit theory of the normal comprehension process.

The device complexity project has two primary goals: the development of a cognitive theory of user-device interaction in including learning and performance models, and the development of a theoretically driven design process that will optimize the relationships between device functionality and ease of learning and other performance factors (Polson & Kieras, 1983, 1984; Polson, Muncher, and Engelbeck 1985). The results of this project should be directly relevant to the design of complex, computer controlled medical equipment. They are currently using word processors to study user-device interactions, but principles underlying use of such devices should generalize to medical equipment.

Both the text comprehension project and the device complexity project involve the development of explicit models of complex cognitive processes; cognitive modeling is a stated goal of both SUMEX and research supported by NIMH.

Several other psychologists have either used or shown an interest in using an early version of the prose comprehension model, including Alan Lesgold of SUMEX's SCP project, who is exporting the system to the LRDC Vax. We have also worked with James Greeno -- another member of the SCP project -- on a project that will integrate this model with models of problem solving developed by Greeno and others at the University of California, Berkeley. Needless to say, all of this interaction has been greatly facilitated by the local and network-wide communication systems supported by SUMEX. The mail system, of course, has also enabled us to maintain professional contacts established at conferences and other meetings, and to share and discuss ideas with these contacts.

D. Progress Summary

The version of the prose comprehension model of 1978 (Kintsch & van Dijk, 1978), which originally was realized as a computer simulation by Miller & Kintsch (1980), has been extended in a major simulation program by Young (1985). Unlike the earlier program, Young includes macroprocessing in her model, and thereby greatly extends the usefulness of the program. It is expected that this program will be widely useful in studies of prose where a detailed theoretical analysis is desired.

The general theory has been reformulated and expanded in van Dijk & Kintsch (1983). This research report of book length presents a general framework for a comprehensive theory of discourse processing. It has been applied to an interesting special case, the question of how children understand and solve word arithmetic problems, by Kintsch & Greeno (1985). A simulation for this model, using INTERLISP, has been supplied in Fletcher (1985).

The device complexity project is in its fourth year. They have developed an explicit model for the knowledge structures involved in the user-device interaction, and they are developing simulation programs. Their preliminary theoretical results are described in Kieras & Polson (1985). They have also completed several experiments evaluating the theory (Polson & Kieras, 1984, 1985; Polson, Muncher, and Engelbeck, 1986) and have shown that number of productions predicts learning time and that number of cycles and working memory operations predicts execution time for a method.

E. List of Relevant Publications

1. Fletcher, R.C.: *Understanding and solving word arithmetic problems: A computer simulation.* Technical Report No. 135, Institute of Cognitive Science, Colorado, 1984.

2. Kieras, D.E. and Polson, P.G.: *The formal analysis of user complexity.* Int. J. Man-Machine Studies, 22, 365-394, 1985.

3. Kintsch, W. and van Dijk, T.A.: *Toward a model of text comprehension and production.* Psychological Rev. 85:363-394, 1978.

4. Kintsch, W. and Greeno, J.G.: *Understanding and solving word arithmetic problems.* Psychological Review, 1985, 92, 109-129.

5. Miller, J.R. and Kintsch, W.: *Readability and recall of short prose passages: A theoretical analysis.* J. Experimental Psychology: Human Learning and Memory 6:335-354, 1980.

6. Polson, P.G. and Kieras, D.E.: *Theoretical foundations of a design process guide for the minimization of user complexity.* Working Paper No. 3, Project on User Complexity, Universities of Arizona and Colorado, June, 1983.

7. Polson, P.G. and Kieras, D.E.: *A formal description of users' knowledge of how to operate a device and user complexity.* Behavior Research Methods, Instrumentation, & Computers, 1984, 16, 249-255.

8. Polson, P.G. and Kieras, D.E.: *A quantitative model of the learning and performance of text editing knowledge.* In Borman, L. and Curtis, B. (Eds.) Proceedings of the CHI 1985 Conference on Human Factors in Computing. New York: Association for Computing Machinery. pp. 207-212, 1985.

9. Polson, P.G. and Jeffries, R.: *Instruction in general problem solving skills: An analysis of four approaches.* In (Eds.) Siegel, J., Chipman, S., and Glaser, R. Thinking and learning skills: Relating instructions to basic research: Vol. 1. Hillsdale, N.J.: Lawrence Erlbaum Associates, pp. 414-455.

10. Polson, P.G., Muncher, E., and Engelbeck, G.: *Test of a common elements theory of transfer.* In Mantei, M. and Orbeton, P. (Eds.) Proceedings of the CHI 1986 Conference on Human Factors in Computing. New York: Association for Computing Machinery. pp. 78-83, 1986.

11. Van Dijk, T.A. and Kintsch, W.: *STRATEGIES OF DISCOURSE COMPREHENSION.* Academic Press, New York, 1983.

12. Young, S.: *A theory and simulation of macrostructure.* Technical Report No. 134, Institute of Cognitive Science, Colorado, 1984.

13. Walker, H.W., Kintsch, W.: *Automatic and strategic aspects of knowledge retrieval.* Cognitive Science, 1985, 9, 261-283.

F. Funding Support Status

7/1/83 – 6/30/84: $56,501

2. Understanding and solving word arithmetic problems
   Walter Kintsch, Professor, University of Colorado
   National Science Foundation
   8/1/83 – 7/31/86: $200,000

3. The Application of Cognitive Complexity Theory to
   the Design of User Interface Architectures
   David Kieras, Associate Professor, University of Michigan
   Peter G. Polson, Professor, University of Colorado
   International Business Machines Corporation
   1/1/85 – 12/31/86: $500,000 (direct+indirect)
   1/1/86 – 12/31/86: $250,000 (direct+indirect)

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Sharing and Interactions with Other SUMEX-AIM Projects

Our primary interaction with the SUMEX community has been the work of the prose comprehension group with the AGE and UNITS projects at SUMEX. Feigenbaum and Nii have visited Colorado, and one of us (Miller) attended the AGE workshop at SUMEX. Both of these meetings have been very valuable in increasing our understanding of how our problems might best be solved by the various systems available at SUMEX. We also hope that our experiments with the AGE and UNITS packages have been helpful to the development of those projects.

We should also mention theoretical and experimental insights that we have received from Alan Lesgold and other members of the SUMEX SCP project. The initial comprehension model (Miller & Kintsch, 1980) has been used by Dr. Lesgold and other researchers at the University of Pittsburgh, as well as researchers at Carnegie-Mellon University, the University of Manitoba, Rockefeller University, and the University of Victoria.

B. Critique of Resource Management

The SUMEX-AIM resource is clearly suitable for the current and future needs of our project. We have found the staff of SUMEX to be cooperative and effective in dealing with special requirements and in responding to our questions. The facilities for communication on the ARPANET have also facilitated collaborative work with investigators throughout the country.

III. RESEARCH PLANS

A. Long Range Projects Goals and Plans

The goal of the prose comprehension project is to develop a computer system capable of the meaningful processing of prose. This work has been generally guided by the prose comprehension model discussed by van Dijk & Kintsch (1983), although our programming efforts have identified necessary clarifications and modifications in that model (Kintsch & Greeno, 1985; Fletcher, 1985; Walker & Kintsch, 1985; Young, 1985). In general, this research has emphasized the importance of knowledge and knowledge-based processes in comprehension. We hope to be able to merge the substantial artificial intelligence research on these systems with psychological interpretations of

E. H. Shortliffe

prose comprehension, resulting in a computational model that is also psychologically respectable.

The primary goal of the device complexity project is the development of a theory of the processes and knowledge structures that are involved in the performance of routine cognitive skills making use of devices like word processors. We plan to model the user-device interaction by representing the user's processes and knowledge as a production system and the device as a nested automata. We are also studying the role of mental models in learning how to use them.

B. Justification and Requirements for Continued SUMEX Use

Both the prose comprehension and the user-computer interaction projects have shifted their actual simulation work from SUMEX to systems at the University of Colorado and the University of Michigan. Both projects use Xerox 1108 systems continuing their work in INTERLISP. However, we consider our continued access to SUMEX critical for the successful continuation of these projects.

Access to SUMEX provides us with continued contact with the SUMEX community, which is especially critical for the prose comprehension project. Knowledge representation languages, e.g. UNITS, and other tools developed by SUMEX are critical for this project. Alternative sources of such software are typically unsatisfactory because the systems have only been developed for use on one project and are typically very poorly documented and less than completely debugged. We hope that our continued membership in the community will be offset by the input that we have been and will continue to provide to various projects: our relationship has been symbiotic, and we look forward to its continuation.

Access to SUMEX's mail facilities are critical for the continued success of these projects. These facilities provide us with the means to interact with colleagues at other universities. Kintsch is currently collaborating with James Greeno, who is at the University of California at Berkeley, and Polson's long-term collaborator, David Kieras, is at the University of Michigan. In addition, our access to the Xerox 1108 (Dandelion) user's community is through SUMEX.

We currently use four computing systems for the VAX 11/780, and three Xerox 1108s, one of which is at the University of Michigan. The VAX is used primarily to collect experimental data designed to evaluate the simulation models and to do necessary statistical analysis.

C. Needs and Plans for Other Computational Resources

SUMEX provides us with two critical needs. The first is communication, which we discussed in the preceding paragraph. The second is technical advice and access to various knowledge representation languages like UNITS.

We envisage our future needs to be communication currently served by the SUMEX 2060 and technical advice and necessary software provided by the SUMEX staff.

D. Recommendations for Future Community and Resource Development

Our future needs are for the SUMEX-AIM resource to act as a communications crossroad and to develop software and provide technical support for user community work stations. We have no preferences as to how such services are provided: either with a communication server on the network, or with the central machine like the current 2060.

We will continue to need access to the SUMEX-AIM 2060 in order to access communication networks and to interact with the SUMEX-AIM staff and community.

If communications and access to the staff are provided through some other mechanism, then we would no longer need access to the 2060.

We would be willing to pay fees for using SUMEX communication resources if required by NIH. However, our willingness is price sensitive. Any charges over $1,000 a year would mean we should communicate with people directly by long-distance telephone.

E. H. Shortliffe

IV.B.3. MENTOR Project

MENTOR Project

Stuart M. Speedie, Ph.D.
School of Pharmacy
University of Maryland

Terrence F. Blaschke, M.D.
Department of Medicine
Division of Clinical Pharmacology
Stanford University

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

The goal of the MENTOR (Medical EvaluatioN of Therapeutic ORders) project is to design and develop an expert system for monitoring drug therapy for hospitalized patients that will provide appropriate advice to physicians concerning the existence and management of adverse drug reactions. The computer as a record-keeping device is becoming increasingly common in hospital-based health care, but much of its potential remains unrealized. Furthermore, this information is provided to the physician in the form of raw data which is often difficult to interpret. The wealth of raw data may effectively hide important information about the patient from the physician. This is particularly true with respect to adverse reactions to drugs which can only be detected by simultaneous examinations of several different types of data including drug data, laboratory tests and clinical signs.

In order to detect and appropriately manage adverse drug reactions, sophisticated medical knowledge and problem solving is required. Expert systems offer the possibility of embedding this expertise in a computer system. Such a system could automatically gather the appropriate information from existing record-keeping systems and continually monitor for the occurrence of adverse drug reactions. Based on a knowledge base of relevant data, it could analyze incoming data and inform physicians when adverse reactions are likely to occur or when they have occurred. The MENTOR project is an attempt to explore the problems associated with the development and implementation of such a system and to implement a prototype of a drug monitoring system in a hospital setting.

B. Medical Relevance and Collaboration

A number of independent studies have confirmed that the incidence of adverse reactions to drugs in hospitalized patients is significant and that they are for the most part preventable. Moreover, such statistics do not include instances of suboptimal drug therapy which may result in increased costs, extended length-of-stay, or ineffective therapy. Data in these areas are sparse, though medical care evaluations carried out as part of hospital quality assurance programs suggest that suboptimal therapy is common.

Other computer systems have been developed to influence physician decision making by monitoring patient data and providing feedback. However, most of these systems suffer from a significant structural shortcoming. This shortcoming involves the evaluation rules that are used to generate feedback. In all cases, these criteria consist of discrete,

independent rules, yet medical decision making is a complex process in which many factors are interrelated. Thus, attempting to represent medical decision-making as a discrete set of independent rules, no matter how complex, is a task that can, at best, result in a first order approximation of the process. This places an inherent limitation on the quality of feedback that can be provided. As a consequence it is extremely difficult to develop feedback that explicitly takes into account all information available on the patient. One might speculate that the lack of widespread acceptance of such systems may be due to the fact that their recommendations are often rejected by physicians. These systems must be made more valid if they are to enjoy widespread acceptance among physicians.

The proposed MENTOR system is designed to address the significant problem of adverse drug reactions by means of a computer-based monitoring and feedback system to influence physician decision-making. It will employ principles of artificial intelligence to create a more valid system for evaluating therapeutic decision-making.

The work in the MENTOR project is intended to be a collaboration between Dr. Blaschke at Stanford and Dr. Speedie at the University of Maryland. Dr. Speedie provides the expertise in the area of artificial intelligence programming. Dr. Blaschke provides the medical expertise. The blend of previous experience, medical knowledge, computer science knowledge and evaluation design expertise they represent is vital to the successful completion of the activities in the MENTOR project.

C. Highlights of Research Progress

The MENTOR project was initiated in December 1983. The project has been funded by the National Center for Health Services Research since January 1, 1985. Initial effort has focused on exploration of the problem of designing the MENTOR system. Work has begun on constructing a system for monitoring potassium in patients with drug therapy that can adversely affect potassium. Antibiotics, dosing in the presence of renal failure, and digoxin dosing have been identified as additional topics of interest.

E. Funding Support

Title: MENTOR: Monitoring Drug Therapy for Hospitalized Patients

Principal Investigators:
>   Terrence F. Blaschke, M.D.
>   Division of Clinical Pharmacology
>   Department of Medicine
>   Stanford University
>
>   Stuart M. Speedie, Ph.D.
>   School of Pharmacy
>   University of Maryland

Funding Agency: National Center for Health Services Research

Grant Identification Number: 1 R18 HS05263

Total Award: January 1, 1985 - December 31, 1988 $485,134 Total
>   Direct Costs

Current Period: January 1, 1986 - December 31, 1986 $182,820 Total
>   Direct Costs

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaborations and Program Dissemination via SUMEX

This project represents a collaboration between faculty at Stanford University Medical Center and the University of Maryland School of Pharmacy in exploring computer-based monitoring of drug therapy. SUMEX, through its communications capabilities, facilitates this collaboration of geographically separated project participants by allowing development work on a central machine resource and file exchange between sites.

B. Sharing and Interactions with Other SUMEX-AIM Projects

Interactions with other SUMEX-AIM projects has been on an informal basis. Personal contacts have been made with individuals working on the ONCOCIN project concerning system development issues. Dr. Perry Miller has also been of assistance by providing software for advisory generation. Given the geographic separation of the investigators, the ability to exchange mail and programs via the SUMEX system as well as communicate with other SUMEX-AIM projects is vital to the success of the project.

C. Critique of Resource Management

To date, the resources of SUMEX have been fully adequate for the needs of this project. The staff have been most helpful with any problems we have had and we are quite satisfied with the current resource management.

III. RESEARCH PLANS

A. Project Goals and Plans

The MENTOR project has the following goals:

1. Implement a prototype computer system to continuously monitor patient drug therapy in a hospital setting. This will be an expert system that will use a modular, frame-oriented form of medical knowledge, a separate inference engine for applying the knowledge to specific situations, and automated collection of data from hospital information systems to produce therapeutic advisories.

2. Select a small number of important and frequently occurring medical settings (e.g., combination therapy with cardiac glycosides and diuretics) that can lead to therapeutic misadventures, construct a comprehensive medical knowledge base necessary to detect these situations using the information typically found in a computerized hospital information system and generate timely advisories intended to alter behavior and avoid preventable drug reactions.

3. Design and begin to implement an evaluation of the impact of the prototype MENTOR system on physicians' therapeutic decision-making as well as on outcome measures related to patient health and costs of care.

1986 will be spent on prototype development in four content areas, design and implementation of the basic knowledge representation and reasoning mechanisms, and preliminary interfacing to existing patient information systems.

B. Justification and Requirements for Continued SUMEX Use

This project needs continued use of the SUMEX facilities for two reasons. First, it

provides access to an environment specifically designed for the- development of AI systems. The MENTOR project focuses on the development of such a system for drug monitoring that will explore some neglected aspects of AI in medicine. This environment is necessary for the timely development of a well-designed and efficient MENTOR system. Second, access to SUMEX is necessary to support the collaborative efforts of geographically separated development teams at Stanford and the University of Maryland. Furthermore, the MENTOR project is predicated on the access to the SUMEX resource free of charge over the next two years. Given the current restrictions on funding, the scope of the project would have to be greatly reduced if there were charges for use of SUMEX.

C. Needs and Plans for Other Computing Resources Beyond SUMEX-AIM

A major long-range goal of the MENTOR project is to implement this system on a independent hardware system of suitable architecture. It is recognized that the full monitoring system will require a large patient data base as well as a sizeable medical knowledge base and must operate on a close to real-time basis. Ultimately, the SUMEX facilities will not be suitable for these applications. Thus we intend to transport the prototype system to a dedicated hardware system that can fully support the the planned system and which can be integrated into the SUMC Hospital Information System. For this purpose a VAX 750 and two XEROX 1186 workstations have been acquired and our development efforts are gradually being transferred to them.

D. Recommendations for Future Community and Resource Development

In the brief time we have been associated with SUMEX, we have been generally pleased with the facilities and services. However, it is clearly evident that the users' almost insatiable demands for CPU cycles and disk space cannot be met by a single central machine. The best strategy would appear to be one of emphasizing powerful workstations or relatively small, multi-user machines linked together in a nationwide network with SUMEX serving as the its central hub. This would give the individual users much more control over the resources available for their needs, yet at the same time allow for the communications among users that have been one of SUMEX's strong points.

For such a network to be successful, further work needs to be done in improving the network capabilities of SUMEX to encourage users at sites other that Stanford. Further work is also needed in the area of personal workstations to link them to such a network. Given the successful completion of this work, it would be reasonable to consider the gradual phase-out of the central SUMEX machine over two or three years and its replacement by an efficient, high-speed communications server.

E. H. Shortliffe

IV.B.4. Rutgers Research Resource

Rutgers Research Resource--Artificial Intelligence in Medicine

Principal Investigators:
Casimir Kulikowski, Sholom Weiss
Rutgers University, New Brunswick, New Jersey

I. SUMMARY OF RESEARCH PROGRAM

A. Goals and Approach

The fundamental objective of the Rutgers Resource is to develop a computer based framework for advancing research in the biomedical sciences and for the application of research results to the solution of important problems in health care. The central concept is to introduce advanced methods of computer science - particularly in artificial intelligence - into specific areas of biomedical inquiry. The computer is used as an integral part of the inquiry process, both for the development and organization of knowledge in a domain and for its utilization in problem solving and in processes of experimentation and theory formation.

An essential part of the resource is directed to methodological problems of knowledge representation and to the development of computer-based systems for acquiring, managing, and improving knowledge bases, and for constructing expert reasoning models in medicine. Equally fundamental are the problems of how best to use knowledge bases and models in processes of interpretation/diagnosis, planning, theory formation, simulation, and effective man-machine communication. These are problems we are studying in the Resource in the context of several system building efforts that address themselves to specific tasks of clinical decision-making and model development and testing.

Resource activities include research projects (collaborative research and core research) training/dissemination projects, and computing services in support of user projects.

B. Medical Relevance and Collaborations

In 1985-86 we continued the development of several versatile systems for building and testing consultation models in biomedicine. The EXPERT system has had many of its capabilities enhanced in the course of collaborative research in the areas of rheumatology, ophthalmology, and clinical pathology.

In *ophthalmology* we have developed a knowledge representation scheme for treatment planning which is both natural and efficient for encoding the strategies for choosing among competing and cooperating treatment plans. This involves a ranking of treatments according to their characteristics and desired effects as well as contraindications. A diagnosis and treatment planning program for ocular herpes was developed using this scheme. Our main collaboration continues to be with Dr. Chandler Dawson of the Proctor Foundation, UCSF.

In *rheumatology*, the model for rheumatological diseases now includes detailed diagnostic criteria for 26 major diseases. The management advice and treatment planning has been developed further. The Resource researchers have developed new representational elements for EXPERT in response to the needs of the rheumatology research Politakis originally developed a coordinated system called SEEK (System for

Empirical Experimentation with Expert Knowledge) which provides interactive
assistance to the human expert in testing, refining and updating a knowledge base
against a data base of trial cases. A generalized version of SEEK, SEEK2, has been
developed during the past year. Dr. Lindberg of the National Library of Medicine, and
Dr. Sharp of the University of Missouri are the project leaders in developing the
rheumatology knowledge base for this effort.

In *clinical pathology* our main collaboration has been with Dr. Robert Galen
(Cleveland Clinic Foundation), with whom we have developed the serum protein
electrophoresis model which is incorporated into an instrument with a scanning
densitometer. This instrument with interpretive reporting capabilities has now been on
the market for over a year, and is located at several hundred clinical sites. We are
making good progress developing a knowledge based system for the interpretation of
CPK/LDH isoenzymes.

In biomedical modeling applications we are experimenting with several prototype
models for giving advice on the interpretation of experimental results in the field of
enzyme kinetics, in conjunction with Dr. David Garfinkel. His PENNZYME program
has been linked to a model in EXPERT, which allows the user to interpret the progress
of the model analysis, and a framework for the design of experiments in this domain
has been formulated.

C. Highlights of Research Progress

Research has continued on problems of representation, inference and control in expert
systems. Emphasis has been placed this year on problems of knowledge base
acquisition, empirical testing and refinement of reasoning (the SEEK2 system). From a
technological point of view, the market availability of the interpretive reporting version
of a scanning densitometer, and the development of models for eye care consultation
that run on microprocessor systems (Apple IIe, IBM-PC) represents an important
achievement for AIM research in showing its practical impact in medical applications.

- Knowledge Base Refinement: SEEK is a system which has been developed
to give interactive advice about rule refinement during the design of an
expert system. The advice takes the form of suggestions for possible
experiments in generalizing and specializing rules in an expert model that
has been specified based on reasoning rules cited by a human expert. Case
experience, in the form of stored cases with known conclusions, is used to
interactively guide the expert in refining the rules of a model. The design
framework of SEEK consists of a tabular model for expressing expert-
modeled rules and a general consultation system for applying a model to
specific cases. This approach has proven particularly valuable in assisting
the expert in domains where the logic for discriminating two diagnoses is
difficult to specify, and we have benefited primarily from experience in
building the consultation system in rheumatology. During the past year, a
newer SEEK2 system has been developed that has enhanced capabilities,
including a more generalized knowledge base and an automatic pilot
capability to proceed with knowledge base refinements. The original work on
SEEK was recognized by the international HUSPI award for medical expert
system research.

- Technology Transfer: Important technology transfer milestones have also
been achieved this year: the instrument interpretation EXPERT program for
serum protein has been widely disseminated, as has the Ocular Herpes
Treatment Program.

E. H. Shortliffe

D. Up-to-Date List of Publications

The following is an update of publications in the Rutgers Resource for the period 1984 and 1985 (only publications not listed in previous SUMEX annual reports are presented here).

1. Apte, C. and Weiss, S., An Approach to Expert Control of Interactive Software Systems, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 7 No. 5, pp. 586-591, (1985).

2. Ginsberg, A., Weiss, S., and Politakis, P., SEEK2: A Generalized Approach to Automatic Knowledge Base Refinement, *Proceedings of the 1985 International Joint Conference on Artificial Intelligence* pp. 367-374 (1985).

3. Ostroff, J., Dawson, C., Kastner, J., Weiss, S., Kulikowski, C., Kern, K., An Expert Advisory System for Primary Eye Care in Developing Countries, *Proceedings 1985 Expert Systems in Government Conference*, pp. 490-495 (1985).

4. Ostroff, J., Dawson, C., Kastner, J., Weiss, S., Kulikowski, C., Kern, K., Preliminary Results from Field Testing of an Expert Advisory System for Primary Eye Care in Developing Countries, Proceedings AI-86 Conference, Long Beach, CA. (1986).

5. Galen, R. and Weiss, S., The CPK/LDH Diagnostic Program, *Proceedings of the 5th International Meeting on Clinical and Laboratory Organization and Management*, Haifa, in press (1985).

6. Weiss, S.M. and Kulikowski, C.A.: *A Practical Guide to Designing Expert Systems*, Rowman and Allanheld, 1984.

7. Kastner, J., Weiss, S., Kulikowski, C., and Dawson, C.: *Therapy Selection in an Expert Medical Consultation System for Ocular Herpes Simplex* Computers in Biology and Medicine, Vol. 14, No. 3, pp. 285-301 (1984).

8. Dawson, C., Kastner, J., Weiss, S., Kulikowski, C.: *A Computer-based Method to Provide Subspecialist Expertise on the Management of Herpes Simplex Infections of the Eye*, Proceedings International Symposium On Herpetic Eye Diseases, Belgium (1984).

9. Galen, R. and Weiss, S.: *Predictive Value Calculator*, American Society of Clinical Pathologists, Clinical Chemistry # CC 84-4 (1984).

10. Kastner, J., Dawson, C., Weiss, S., Kern, K., Kulikowski, C.: *An Expert Consultation System for Frontline Health Workers in Primary Eye Care*, Journal of Medical Systems, Vol. 8, No. 5 (1984).

Indicate by an asterisk (*) that the resource was given credit.

E. Funding Support

E. H. Shortliffe   156

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. *Medical Collaborations and Dissemination*

The SUMEX-AIM facility provides a backup node where some of our medical collaborators can access programs developed at Rutgers. The bulk of the medical collaborative work outlined in I.B. above is centered at the Rutgers facility (the Rutgers-AIM node).

Dissemination activities continue to be an important responsibility of the Rutgers Resource within the AIM community. The following activities took place in the last year:

1. AIM Workshop (1985):

   Organized by Dr. Kingsland of the National Library of Medicine. It consisted of a series of presentations on AIM research and related work by members of the AIM community.

2. 1984 Hawaii International Conference On Systems Sciences:

   Dr. Weiss presented a paper on the expert system for front-line health workers, and Dr. Kulikowski chaired a session on knowledge based medical systems.

B. *National AIM Projects at Rutgers*

The national AIM projects, approved by the AIM Executive Committee, that are associated with the Rutgers-AIM node are the following:

1. Attending Project, directed by Dr. Perry Miller of the Yale Medical Center, is doing much of the research on critiquing a physician's plan of management at Rutgers.

2. Medical Knowledge Representation project, headed by Dr. Chandrasekaran from Ohio State University, is doing most of its research on the Rutgers system.

3. Biomedical Modeling, by Dr. Garfinkel from the University of Pennsylvania.

4. INTERNIST/CADUCEUS project, headed by Dr. Myers and Dr. Pople from the University of Pittsburgh, has been using the Rutgers Resource as a backup system for development and experimentation.

C. *Critique of SUMEX-AIM Resource Management*

Rutgers is currently using the SUMEX DEC-20 system primarily for communication with other researchers in the AIM community and with SUMEX staff, and also for backup computing in demonstrations, conferences and site visits. Our usage is currently running at less than 50 connect hours per year at SUMEX, with an overall connect/CPU ratio of about 30.

Rutgers is beginning to place more emphasis on the use of personal computers, and on network support needed to make these effective. SUMEX has been of significant help in their developmental efforts in networking workstation software.

E. H. Shortliffe

III. RESEARCH PLANS

A. *Project Goals and Plans*

We are planning to continue along the main lines of research that we have established in the Resource to date. Our medical collaborations will continue with emphasis on development of expert consultation systems in rheumatology, ophthalmology and clinical pathology. The basic AI issues of representation, inference and planning will continue to receive attention. Our core work will continue with emphasis on further development of generalized expert system frameworks and also on AI studies in representations and problems of knowledge and expertise acquisition. We propose to work on a number of technology transfer experiments to micro processing that will be affordable by our biomedical research and clinical collaborators. We also plan to continue our participation in AIM dissemination and training activities as well as our contribution -- via the Rutgers computers -- to the shared computing facilities of the national AIM network.

B. *Justification and Requirements for Continued SUMEX Use*

Continued access to SUMEX is needed for:

1. Backup for demos, and so forth.

2. Programs developed to serve the National AIM Community should be runnable on both facilities.

3. There should be joint development activities between the staffs at Rutgers and SUMEX in order to ensure portability, share the load, and provide a wider variety of inputs for developments.

C. *Needs and Plans for Other Computing Resources Beyond SUMEX-AIM*

Our computing needs are based on a centralized computing resource accessible to distant users, and local workstations. We will continue to use SUMEX for backup purposes.

D. *Recommendations for Future Community and Resource Development*

Use of personal computers and workstations is continuing to grow in the AIM community. We find that the biggest challenge is supporting these systems. Although some central computing will continue to be needed for communication and coordination, we believe that over the next few years all AIM research projects and even individual collaborators will come to have their own hardware. However, many of these community members (particularly the collaborators) will not be in a position to support hardware or software on their own. We would certainly expect SUMEX to continue to provide expert advice in this area. However, we believe it would be helpful for SUMEX to have a formal program to support smaller computers in the field. We envision this as including at least the following items:

- A central source of information on hardware and software that is likely to be of interest to the AIM community. SUMEX might want to become a distribution point for certain of this software, and even help coordinate quantity purchase of hardware if this proves useful.

- Assistance in support of hardware and software in the field. Depending upon the hardware involved, this might involve advice over the telephone or actual board-swapping by mail.

IV.B.5. SOLVER Project

SOLVER: Problem Solving Expertise

Dr. P. E. Johnson
Center for Research in Human Learning
University of Minnesota

Dr. James R. Slagle
Department of Computer Science
University of Minnesota

Dr. W. B. Thompson
Department of Computer Science
University of Minnesota

I.   SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

This project focuses upon the development of strategies for discovering and documenting the knowledge and skill of expert problem solvers. In the last several years, considerable progress has been made in synthesizing the expertise required for solving extremely complex problems. Computer programs exist with competency comparable to human experts in diverse areas ranging from the analysis of mass spectrograms and nuclear magnetic resonance (Dendral) to the diagnosis of certain infectious diseases (Mycin).

Design of an expert system for a particular task domain usually involves the interaction of two distinct groups of individuals, "knowledge engineers," who are primarily concerned with the specification and implementation of formal problem solving techniques, and "experts" (in the relevant problem area) who provide factual and heuristic information of use for the problem solving task under consideration. Typically, the knowledge engineer consults with one or more experts and decides on a particular representational structure and inference strategy. Next, "units" of factual information are specified; that is, properties of the problem domain are decomposed into a set of manageable elements suitable for processing by the inference operations. Once this organization has been established, major efforts are required to refine representations and acquire factual knowledge organized in an appropriate form. Substantial research problems exist in developing more effective representations, improving the inference process, and in finding better means of acquiring information from either experts or the problem area itself.

Programs currently exist for empirical investigation of some of these questions for a particular problem domain (e.g. AGE, UNITS, RLL). These tools allow the investigation of alternate organizations, inference strategies, and rule bases in an efficient manner. What is still lacking, however, is a theoretical framework capable of reducing dependence on the expert's intuition or on near exhaustive testing of possible organizations. Despite their successes, there seems to be a consensus that expert systems could be better than they are. Most expert systems embody only the limited amount of expertise that individuals are able to report in a particular, constrained language (e.g. production rules). If current systems are approximately as good as human experts, given that they represent only a portion of what individual human experts know, then

improvement in the "knowledge capturing" process should lead to systems with considerably better performance.

In order to obtain a broad view of the nature of human expertise, the SOLVER project includes studies in a variety of complex problem solving domains in addition to medicine. These include law, auditing, business management, plant pathology, and expert system design. We have observed that despite the apparent dissimilarities in these problem solving areas, there is reason to believe that there are underlying principles of expertise which apply broadly. Our project seeks to investigate these principles and to create tools to make use of that knowledge in practical expert systems.

B. Medical Relevance and Collaboration

Much of our research has been and will continue to be directly focused on medical AI problems. Galen, our experimental expert system in pediatric cardiology, is achieving expert levels of performance. Dr. Connelly is leading a project that is developing an expert system based platelet transfusion therapy monitoring program. Dr. Spackman is completing a doctoral thesis on the automated acquisition of rule knowledge in medical microbiology.

Some of our research has focused on problems in diagnostic reasoning and expertise in domains other than medicine. However, our experience indicates that principles of expertise and relevant knowledge engineering tools can cut across task domains. Galen is demonstrably a useful expert system implementation tool designed in the medical diagnostic task domain. Developments from our work in other domains affecting problems such as automated knowledge acquisition through rule induction and reasoning by analogy will have medical relevance.

Collaboration with Dr. James Moller in the Department of Pediatrics, Dr. Donald Connelly in the Department of Laboratory Medicine, at the University of Minnesota. Dr. Connelly has become a SUMEX user and is teaching a course in medical informatics. He has also initiated a project to create an expert system in platelet transfusion therapy.

Collaboration with Dr. Eugene Rich and Dr. Terry Crowson at St. Paul Ramsey Medical Center.

Dr. Kent Spackman is a post-doctoral fellow in medical informatics who is completing a Ph.D. thesis in Artificial Intelligence. Dr. Spackman is a resident at the University of Minnesota Hospitals and collaborates with the SOLVER project.

Recent research by Prof. Slagle has addressed the issue of economical information acquisition in medical diagnostic expert systems. A demonstration system has been implemented studying cases from the Program on the Surgical Control of the Hyperlipidemias. This expert system evaluates the results of treadmill ECG examinations.

C. Highlights of Research Progress

Accomplishments of This Past Year --

Dr. Slagle has published the results of his work on the design of the Merit system. Merit enables an expert system to direct the acquisition of information by finding questions with a high ratio of probable importance to difficulty. Merit has been incorporated in a general purpose expert system shell called GENIE (Generalized Inference Engine). In addition, Dr. Slagle has described an expert system that can be used to support the data collection and evaluation problems in a large clinical study: the Program on the Surgical Control of Hyperlipidemias (POSCH). POSCH is a long term randomized clinical trial with 838 post-myocardial infarction patients.

Dr. Connelly has supervised the implementation of two medical expert systems. One deals with *detection of deviations in time series by the human observer* and is an expert system to assist observers in the monitoring of laboratory data. The other is a *knowledge based system for improving transfusion practice.* The object of this system is to promote nearly optimal usage of platelet concentrates in a tertiary care setting through the use of an expert system.

Dr. Spackman has developed an algorithm useful for inductive inference in generating classification rules from empirical examples. The algorithm has been tested with data from several medical domains, including microbiology, endocrinology and pediatric cardiology.

Dr. Johnson has been engaged in continuing studies of medical expertise. Current research has included the study of memory processes of diagnosticians and the role of model-based reasoning in diagnosis. In addition, a project studying the role of context in analogical reasoning is under way, and a prototype inference engine using context directed reasoning by analogy has been implemented.

Drs. Thompson and Johnson have also undertaken a study of explanation facilities in expert systems. Explanation is related to the line of reasoning of the system and the needs of the users.

Research in Progress --

Application of Galen in different problem domains

One project to test the extensibility of Galen into other domains has recently been completed in an area of accounting (internal audit controls). The objective of the study was to formulate and test a model of the processes employed by audit managers and partners in reviewing and evaluating internal accounting controls. In collaboration with an international accounting firm, a rule base was developed and successfully tested against the actions of expert auditors.

Another project explored the extension of the Galen architecture into a problem in agriculture. The main objective of this research was to determine whether the basic postulates about expert reasoning made in Galen would apply successfully in another diagnostic domain. The problem domain chosen for this purpose was plant pathology. The system is currently going through evaluation and fine tuning to bring it up to an expert performance level. This system may be useful in the Extension Service at the University of Minnesota, which provides diagnostic information to farmers over the phone lines.

Inference engine mechanisms

In another area of research, a small rule based program called PROTEUS was recently completed for the problem of computer hardware diagnosis. Several hardware diagnostic programs are currently based on structures that are essentially equivalent to decision trees. As such a program's competence becomes larger, however, the program becomes increasingly difficult to modify. The purpose of the PROTEUS project was to investigate how a rule based system could be used to represent the same diagnostic knowledge as a decision tree and yet be easier to understand and modify due to the modular nature of production rules. PROTEUS represents the state of the problem being solved as a tree of symbols. For example, some portions of the tree represent lines of questioning that should be followed to get information about the problem. Other portions represent facts that have been obtained as a result of previous questions. Still others represent conclusions that were drawn from the these facts. PROTEUS works by using rewrite rules to transform the tree. A rewrite rule recognizes a subtree that matches the pattern in its antecedent and either replaces or modifies this subtree

E. H. Shortliffe

with a new tree from its consequent. This single general mechanism is used to ask questions, draw conclusions from the answers, and modify the line of questioning accordingly.

Merit system for question selection

Another direction in our current research has focused on developing a generalized inference engine as an extension and improvement of the ideas found in BATTLE, an expert consultant system for weapon to target allocation. The BATTLE inference engine models the relevant considerations for allocating a set of weapons to a set of targets. The system incorporates variants of several important techniques of artificial intelligence and makes the first use of the Merit system for question selection. The Merit system enables the system to direct the acquisition of information by finding questions with a high ratio of probable importance to difficulty. In the first phase of the allocation process, the system uses a computation network to determine the effectiveness of each individual weapon against each prospective target. The network, built by a domain expert in advance, allows reasoning with logical, Bayesian, and expert-defined operators. After the calculation of individual effectiveness values, a portion of the allocation tree is built to determine good allocation plans for the set of weapons. The individual effectiveness values are used to direct the traversal and pruning of the allocation tree.

This approach has been extended in the GENIE (GENeralized Inference Engine) system. GENIE was designed to provide greater representational power, simplified knowledge specification, and smaller storage requirements. GENIE can manipulate elementary concepts involving several objects at once and can represent the subset-superset relationships of the types of objects. The elementary concepts may have values taken from any well-defined data type. Moreover, representational structures such as nodes and edges are shared whenever possible, and some redundancies of the previous system are reduced, resulting in an increase of storage efficiency in GENIE. The current version of GENIE includes the use of the Merit questioning system and a menu-driven interface. GENIE not only represents a significant achievement in research on inference engines, but also provides as excellent tool for future research on Merit.

Detection of deviations in time series by the human observer

Surveillance and early detection of deviation from a homeostatic state are goals common to health care programs for the apparently healthy as well as for groups of patients known to have or have had specific diseases. Automated approaches to detecting deviations have the advantage of being reliably applied, traceable, consistent in outcome, and conserving of professional resources. However, these techniques may be relatively insensitive, require a moderately large number of observations before they become functional, and are not necessarily tuned to the detection of the specific type of deviation of clinical significance. The human observer using graphical trend displays coupled with specific domain knowledge may be a more sensitive and reliable detector of deviations at an earlier time in the course of clinical observations. But the human observer may be inconsistent, unavailable, distracted or expensive. By embodying the human rules of deviation detection in an expert system, perhaps the advantages of the human expert can be gained while at the same time the disadvantages are eliminated. A study has been done to test these hypotheses. Time trend graphs representing monthly monitoring of serum carcinoembryonic antigen (CEA) levels in simulated patients with surgically-removed breast cancer were presented to six human observers and a time series analysis based on a homeostatic model. Three of the human observers described their rationale in assigning a level of suspicion regarding the presence of an important deviation as the observation points on the graphs were serially revealed to them. The verbalizations were analyzed to develop rules for an expert system. Based on a small

number of cases, the false positive rate of all three approaches (i.e., analytic method, observers, and expert system) appeared comparable. Observers detected deviations earlier than the analytic method in 13 of 39 instances and tied in 15. The expert system tied the observers in 19 of 38 instances and surpassed them seven times. This preliminary study suggests that human observers using time trend graphs can supplement analytic techniques in the early detection of deviations in biologic time series and that the human skill of graph reading can at least partially be extracted and embodied in an expert system.

Knowledge based system for improving transfusion practice

The goal of this project was to promote more nearly optimal usage of platelet concentrates in a tertiary care setting. Through the use of an expert system operating as an information gathering aid to blood bank personnel, professional knowledge and judgment will consistently, automatically and economically be brought to bear on each request for platelet concentrates. The approach used should lead to more consistent and complete professional overview and coordination of blood product use, the development of general and specific guidelines for appropriate use of platelet concentrates, decreased blood bank administrative costs related to blood product use review, and decreased costs to patients. Finally, the project will demonstrate an approach to medical care quality assurance that more realistically deals with the complexity, variability and breadth of clinical problems than is possible with conventional criteria list techniques.

The specific aims of the research are:

1. to develop a set of transfusion practice guidelines for platelet concentrates and to employ these guidelines in an active tertiary care setting, to develop a supplementary guideline set in the form of a knowledge base allowing disease-specific, therapy-specific, trend and other contingencies to be taken into account when considering the appropriateness of transfusion, to develop an inexpensive computer-based monitoring system using expert system techniques to semi-automatically evaluate adherence to these transfusion guidelines,

2. to develop an automatically maintained platelet use database for providing physicians with practice pattern feedback, and to demonstrate the effect of this approach on platelet concentrate use and administrative costs.

At the present time, the platelet transfusion knowledge base acquired so far has been incorporated into an expert system implemented in TLC LISP running on an HP Vectra microcomputer. Automatic acquisition of laboratory data has been accomplished. Work is continuing in knowledge acquisition and computer to computer communication capability.

Inductive Inference for Medical Diagnosis

An inductive inference algorithm has been developed for generating classification rules from empiric examples. Such rules may be useful for the conceptual analysis of data and for the construction and refinement of knowledge bases for expert systems. This research addresses some deficiencies of currently available methods for deriving classification rules from empiric data, and explores the use of unate Boolean functions as a language for logical classification rules.

The goals of this research are to find computational methods for deriving comprehensible classification rules from empiric data, and to evaluate the performance of these methods in producing formal criteria for classification. Further, this research attempts to find and develop methods for creating rules (concept descriptions) which use a format known as "criteria tables" or "counting rules."

E. H. Shortliffe

Such criteria tables may be viewed more abstractly as Boolean functions. In addition, they fall into a category known as unate Boolean functions, which may be defined as the set of logical expressions that can be formed using the connectives AND and OR, without the NOT connective. The algorithm that has been developed takes advantage of some of the interesting and useful properties of unate functions.

Currently available methods in machine learning, pattern recognition, and statistical discriminant analysis have several deficiencies that limit their usefulness in generating descriptions of concepts or in building expert systems. The pattern recognition and discriminant analysis methods fail to produce human-oriented, comprehensible results, and they do not take into account background knowledge about the problem domain. Machine learning algorithms (such as Michalski's AQ or Quinlan's ID3) produce logic-based results, but they have other drawbacks. For example, some cannot deal with inconsistencies, incomplete data or noise in the training examples; most produce discriminants which incorrectly overfit the data. Because they use languages with essentially unlimited expressiveness, they place no bound on the complexity of the derived rules, leading to complicated rules and a high degree of instability of the rules with varying training sets. Most provide for little or no incorporation of background knowledge.

To address these deficiencies, a logic-based approach to inductive inference is combined with background knowledge in an algorithm that allows appropriate handling of inconsistencies and noise, producing concept descriptions that can be expressed as unate Boolean functions or may be further simplified to criteria tables. A Boolean function is unate if no variable appears both complemented and uncomplemented in its minimum sum-of-products (disjunctive normal form) representation. From an intuitive point of view, unateness means that an individual piece of evidence (a variable or proposition) either contributes for or against the conclusion of the rule, but never reverses its influence when combined with other evidence.

The algorithm that has been developed accepts both binary and non-binary variables as input, and then creates propositional variables from the non-binary ones using techniques related to Michalski's AQ algorithm. It also accepts user-supplied domain and task knowledge which it uses to derive rules that more closely reflect the background knowledge and goals of humans. For example, the diagnostic performance goal of the rule can be specified. Possible goals may be 1) to minimize the expected mis-classification error (or cost), 2) to maximize the sensitivity, or 3) to maximize the specificity of the rule. The algorithm then handles conflicting examples (examples with identical features but belonging to different classes) by finding the unate function that best satisfies the given goal of the rule to be derived.

Criteria tables derived by the algorithm may contain an arbitrarily large number of categories of findings (e.g. major, minor, required, exclusions, etc.) but when the number of categories becomes too large, the comprehensibility of the result diminishes. Methods for creating a hierarchy or network of intermediate rules may help to improve comprehensibility in such cases, and developing such methods is a long-term goal of this research.

The algorithm has been tested with data from several medical domains, including microbiology, endocrinology and pediatric cardiology. The results from microbiology show that unate functions are adequate for expressing rules for differentiating groups of Enterobacteriaceae according to a panel of biochemical tests. The results from endocrinology indicate that machine learning adds new information above and beyond what is derivable using linear discriminant analysis. The results from pediatric cardiology reveal that most of the logical competitor sets of the Galen system may be differentiated equally well in thousands of different ways (i.e., with thousands of different criteria tables), adding confirmation to the impression that each expert diagnostician has his own individual strategies that appear to be equally valid.

An unanswered question is whether unate concepts are more natural or more easy to learn than conjunctive concepts or disjunctive concepts. The problem becomes more difficult to answer when one considers the tendency to use general rules of thumb modified by exceptions, because the empirical exceptions may violate the unateness condition and thus make it uncertain whether a concept is inherently unate or whether the exceptions necessarily make the concept disjunctive. Nonetheless, the comprehensibility of unate functions seems to be related to their correspondence to psychological models of the way people categorize.

D. List of Relevant Publications

1. Bailey, A. et al.: "Auditing, artificial intelligence, and expert systems," *Decision Support Systems: Theory and application*, A.B. Whinston (ed.), North-Holland Publications, 1985.

2. Connelly, D. and Johnson, P.E.: *Medical problem solving*. Human Pathology, 11(5):412-419, 1980.

3. Elstein, A., Gorry, A., Johnson, P. and Kassirer, J.: *Proposed Research Efforts*. In D.C. Connelly, E. Benson and D. Burke (Eds.), CLINICAL DECISION MAKING AND LABORATORY USE. University of Minnesota Press, 1982, pp. 327-334.

4. Feltovich, P.J.: *Knowledge based components of expertise in medical diagnosis*. Learning Research and Development Center Technical Report PDS-2, University of Pittsburgh, September, 1981.

5. Feltovich, P.J., Johnson, P.E., Moller, J.H. and Swanson, D.B.: *The Role and Development of Medical Knowledge in Diagnostic Expertise*. In W. Clancey and E.H. Shortliffe (Eds.), READINGS IN MEDICAL AI, Addison-Wesley, 1984, pp. 275-319.

6. Johnson, P.E.: *Problem Solving*. In ENCYCLOPEDIA OF SCIENCE AND TECHNOLOGY, McGraw-Hill, 1985.

7. Johnson, P.E., Moen, J.B., and Thompson, W.B.: *Garden Path Errors in Medical Diagnosis*. IN Bloc, L. and Coombs, M.J. (Eds.), COMPUTER EXPERT SYSTEMS, Springer-Verlag (in press).

8. Johnson, P.E.: *Cognitive Models of Medical Problem Solvers*. IN D.C. Connelly, E. Benson, D. Burke (Eds.), CLINICAL DECISION MAKING AND LABORATORY USE. University of Minnesota Press, 1982, pp. 39-51.

9. Johnson, P.E.: *What kind of expert should a system be?* Journal of Medicine and Philosophy, 8:77-97, 1983.

10. Johnson, P.E.: *The Expert Mind: A New Challenge for the Information Scientist* IN Th. M. A. Bemelmans (Ed.), INFORMATION SYSTEM DEVELOPMENT FOR ORGANIZATIONAL EFFECTIVENESS, Elsevier Science Publishers B. V. (North-Holland), 1984.

11. Johnson, P.E., Severance, D.G. and Feltovich, P.J.: *Design of decision support systems in medicine: Rationale and principles from the analysis of physician expertise*. Proc. Twelfth Hawaii International Conference on System Science, Western Periodicals Co. 3:105-118, 1979.

12. Johnson, P.E., Duran, A., Hassebrock, F., Moller, J., Prietula, M., Feltovich, P. and Swanson, D.: *Expertise and error in diagnostic reasoning*. Cognitive Science 5:235-283, 1981.

13. Johnson, P.E. and Hassebrock, F.: *Validating Computer Simulation Models of Expert Reasoning.* IN R. Trappl (Ed.), CYBERNETICS AND SYSTEMS RESEARCH. North-Holland Publishing Co., 1982.

14. Johnson, P.E. and Thompson, W.B.: *Strolling down the garden path: Detection and recovery from error in expert problem solving.* Proc. Seventh IJCAI, Vancouver, B.C., August, 1981, pp. 214-217.

15. Johnson, P.E., Hassebrock, F. and Moller, J.H.: *Multimethod study of clinical judgement.* Organizational Behavior and Human Performance 30:201-230, 1982.

16. Moller, J.H., Bass, G.M., Jr. and Johnson, P.E.: *New techniques in the construction of patient management problems.* Medical Education 15:150-153, 1981.

17. Sedlmeyer, R.L., Thompson, W.B. and Johnson, P.E.: *Knowledge-based fault localization in debugging.* The Journal of Systems and Software, vol. 3, no. 4 (Dec 83) pp. 301-307, Elsevier.

18. Sedlmeyer, R.L., Thompson, W.B. and Johnson, P.E.: *Diagnostic reasoning in software fault localization.* Proc. Eighth IJCAI, Karlsruhe, West Germany, August, 1983.

19. Slagle, J. et al.: "Expert systems in medical studies -- A new twist," *Proceedings of the Conference on Applications of Artificial Intelligence,* SPIE, 1986.

20. Slagle, J.: "An expert system for a resource allocation problem," *Communications of the ACM,* September, 1985.

21. Slagle, J.: "Alpha Beta Pruning," *Encyclopedia of artificial intelligence,* S. Shapiro (ed.), John Wiley and Sons, Inc., New York, 1985.

22. Slagle, J. and Gaynor, M.: "An intelligent control strategy for computer consultation," *IEEE Trans. on Pattern Analysis and Machine Intelligence,* March, 1984.

23. Smith, K.A., Farm, B., Johnson, P.E.: *Surface: A prototype expert system for selecting surface analysis techniques.* Proceedings of IEEE Conference on Computers and Comm., 1985.

24. Swanson, D.B.: *Computer simulation of expert problem solving in medical diagnosis.* Unpublished Ph.D. dissertation, University of Minnesota, 1978.

25. Swanson, D.B., Feltovich, P.J. and Johnson, P.E.: *Psychological Analysis of Physician Expertise: Implications for The Design of Decision Support Systems.* In D.B. Shires and H. Wold (Eds.), MEDINFO77, North-Holland Publishing Co., Amsterdam, 1977, pp. 161-164.

26. Thompson, W.B., Johnson, P.E. and Moen, J.B.: *Recognition-based diagnostic reasoning.* Proc. Eighth IJCAI, Karlsruhe, West Germany, August, 1983.

E. Funding and Support

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaborations and Program Dissemination via SUMEX

Work in medical diagnosis is carried out with the cooperation of faculty and students in the University of Minnesota Medical School and St. Paul Ramsey Medical Center.

The Galen system is now available from the University of Minnesota as an unsupported research tool for the study of recognition based reasoning systems.

B. Sharing and Interactions with Other SUMEX-AIM Projects

Dr. Paul Johnson participated in a panel discussion with Dr. Harry Pople at the International Conference on Information Systems, Indianapolis, Ind., Dec. 1985. The panel addressed issues in the design and implementation of practical expert systems and expert system research.

C. Critique of Resource Management (None)

III. RESEARCH PLANS

A. Project Goals and Plans

Model based diagnostic reasoning

Since human experts are notoriously poor at describing their own knowledge, one of our planned research objectives is to investigate problem solving tasks through which experts can reveal criteria for initiating specific hypotheses and methods for investigating those hypotheses. We have recently begun preliminary work into how symbolic models of a physical system may be used to acquire knowledge for diagnosing the cause of its observed faults. Previous work at Minnesota on both the Galen and PROTEUS projects (see above) suggests that a model based approach could be useful in acquiring diagnostic knowledge from experts as well as providing a framework for building an expert system that puts this knowledge to use.

We plan to investigate a model seen as a network of component frames that are connected by bidirectional communication links. Each frame represents a real or imaginary component of the system that is being studied. Each link represents a possible dependency between the states of the two components that it connects. For example, a model of a computer might include a memory frame, an arithmetic unit frame, and a bus frame. The arithmetic unit frame might be linked to the bus frame, and the bus frame linked to the memory frame. Symbolically represented information can flow back and forth between frames along links. Other frames that are not part of the computer itself may be included as well. For instance, it may be useful to imagine that all frames are linked to a general environment frame that describes conditions under which the computer is operating.

A component frame has an internal state that is represented symbolically within it. It also has a set of rules that specify relationships between the internal state and the information flowing in and out through the communications links. These rules operate bidirectionally. If a symptom is presented to a frame through one of its links, the frame's rules deduce an internal state that could plausibly cause the symptom. Alternatively, if the internal state changes so that a symptom would be expected, the rules work in the other direction and deduce a symptom that is sent out along a link. The model therefore functions much like a constraint network.

Once the physical model is built, it can be used as the heart of a diagnostic expert system. To use the model in diagnosis, descriptions of symptoms are placed in the internal state of its environment frame. Rules in this frame are applied to suggest changes in the internal states of other frames by sending results along their common communication links. These changes propagate through the model in the same way. Finally the rules in some frame deduce possible faults that could explain the observed symptoms. Each possible fault also causes conjectured symptoms to be deduced. Checks for the presence or absence of these symptoms could be used to confirm or disconfirm the fault. Results of these checks could then be fed back into the model so that incorrect conjectured faults can be removed and correct ones can be strengthened.

One question addressed in this work is what we call the *"inversion problem"* . While it is relatively easy to obtain knowledge from experts about what faults cause what symptoms, it is more difficult to obtain knowledge about what symptoms allow the diagnosis of faults. It is tempting to suppose that the first kind of knowledge, determining symptoms given faults, could be obtained and *inverted* to obtain the more diagnostically useful knowledge about determining faults given symptoms. This is difficult because one symptom may be caused by several different faults, so new knowledge must be provided to decide between them. Such information includes, but is not limited to: What level of detail and abstraction is appropriate for diagnostic reasoning about a system? What faults are the most plausible causes of an individual symptom? What procedures or tests can be executed to investigate a conjectured fault? How can the results suggest what to do next?

Context directed analogical reasoning

A second area of planned work is the investigation of reasoning by analogy within the context of law. Lawyers reason using legal rules, and they reason about the formation and foundation of the rules themselves. While current expert systems technology supports rule-based reasoning when the underlying rules are fixed, the processes involved in reasoning about rules is less well understood. An important feature of legal reasoning is the use of precedents. Precedents are used to discover legal rules and to formulate legal arguments about such rules. We plan to develop a computer model of reasoning with precedents that will be used to investigate the role of "context" in selection and use of precedents. The context selection mechanism is hypothesized to be represented by a system like the Galen system. The process of reasoning with precedents is an elaboration of the model proposed by Winston. The goal of this project is to develop a system that can simulate the reasoning processes of an attorney in selecting relevant precedents from a "knowledge-base" of case law.

Machine learning

A third area of planned work is machine learning. One research project currently underway is developing methods for the automated generation of comprehensible decision rules from empiric data, with emphasis on logic-based knowledge representation formats and on problems drawn from the domain of medicine. This work builds on some of the machine learning methodologies developed at the University of Illinois by R. S. Michalski and others.

This work addresses two shortcomings of previous work on induction of classification rules. These are: first, lack of comprehensibility of the induced rules; and second, lack of flexibility in specifying the diagnostic performance (sensitivity, specificity, or efficiency) desired for the rules that are to be derived.

Comprehensibility of the derived rules or descriptions can be enhanced by imposing restrictions on the format that the rules may take. For example, the restriction of rules to a unate Boolean function format follows the induction of rules that can often be simplified to a *"criteria table"* type of representation. The type of diagnostic performance a rule must have will depend on its purpose, and specifying the purpose may allow inductive inference algorithms to trade off small decrements in diagnostic performance for large increments in comprehensibility, or to increase their robustness in the face of noisy or uncertain data.

Successful development of these techniques will lead to enhanced capabilities for deriving rule bases for expert classification systems from empiric data, and will provide new methods for the conceptual analysis of data.

Preliminary results have been obtained for the problem of deriving rules for the identification of bacteria based on their biochemical profiles in the medical microbiology lab. Other problem domains under investigation are the analysis and interpretation of endocrine laboratory tests, and the induction of rules for the diagnosis of congenital heart disease, for comparison with the rules used in Galen.

Related research is also under way to determine methods of augmenting the process of acquiring the knowledge used to build rule bases for expert classification systems. The premises of this research are that efficient knowledge acquisition depends on the choice of representation, and that in classification domains the concept of *inconsistency* plays an important role. The objective of this research is to design, implement, and test a representation of the beliefs of an *"ideally rational agent"* that is capable of explicitly representing, automatically detecting, and using logical inconsistencies in the agent's set of beliefs. These inconsistencies form the basis for rules that generate class hypotheses, discriminate among them, reject invalid hypotheses, detect noisy or bad input data, and that can direct the search for unknown facts by a questioning facility.

We plan to continue development of a representational formalism for classification knowledge bases that enables inconsistencies to be manipulated as described above. This system is based on an existing nonmonotonic logic belief called *autoepistemic logic* (as defined by Robert Moore) and modified to include the principles of *relevance logic* (as defined by Anderson and Belnap). Having thus defined a representation for the knowledge base, we plan to develop a method for instantiating its concepts within a given application domain. Preliminary research results suggest that an efficient process for learning a rule base consists of first adding the expectations for each class into the logic base and then introducing test cases to the system one at a time. Because the system can automatically detect inconsistencies and determine their type and significance, it can suggest both that a rule is required and what its type should be. We expect the resulting rule base to be more focused and efficient, and to more closely duplicate the lines of reasoning of domain experts. The rule acquisitions tools will be tested in two different applications domains. The resulting expert systems will be evaluated for correctness of classification and similarity of their lines of reasoning with those of human experts.

Intelligent query generation

A fourth area of planned work will be on improving and extending the Merit control strategy described earlier. This is a new computer strategy for deciding what piece of evidence to look for at each stage of the decision process. Every proposition of interest is assigned a weight (called a merit value) proportional to its ability to alter the

decision. At each stage of the consultation, the user is asked the question having the highest merit value. An efficient algorithm has been devised to find the proposition have the highest merit. After the relevant evidence has been gathered, the algorithm updates merit values to reflect the new information. The advantage of this approach is that in a time-critical task, the user is assured that the most critical questions are asked first. Therefore, if questioning must be prematurely terminated, the time has been used to maximum advantage.

The current Merit system, as implemented in BATTLE, is theoretically based on an infinitesimal change to the value of a top node caused by an infinitesimal change to the value of a node on a tree. Merit is efficient, so that it can handle a BATTLE knowledge base with about 700 nodes on the inference network and a data base with about 5,000 data items. Although Merit has worked well experimentally, it is desirable to extend it to handle six situations.

1. Merit should be defined and used relative to any set of nodes, not just one top node.

2. Merit should handle a potential finite change to the value.

3. Merit should be based on a directed acyclic graph, not just a tree.

4. Merit should handle the propagation of distributions, not just single values.

5. Merit should be made practical for a larger knowledge base (7,000 nodes on the inference network).

6. Merit should be made practical for a larger data base (50,000 data items).

Finally, we are looking at ways in which expert systems interact with their environments. Within the context of solving classification problems, we intend to investigate two related aspects of this interaction: *"query generation"* and *explanation* . Some simple expert systems proceed from a fixed set of input data to an evaluation of that data. For most problem domains, however, the space of possibly relevant information is large, and some or all of this information may have costs associated with its acquisition. Thus, computational and other costs can be reduced by some mechanism that intelligently selects appropriate queries designed to solicit information that is relevant and cost effective in terms of the problem being solved. Expert systems for complex problem domains must also be able to generate explanations for their actions. Unless the system operates entirely autonomously, users must be apprised of the rational for system actions. There is a particular need for explanations tailored for system users rather than system designers.

In its general form, the query generation problem arises any time an expert system requires outside information. This information typically can be provided my one of three sources: the *user, sensors,* or *"other systems"* . Most current systems get information about a particular problem by directly querying a user. Data entry is slow and user patience limited, so it is important that the number of questions asked be restricted and the particular questions asked seem reasonable to the user. Some expert systems have direct access to sensors. The selection of sensor data becomes important if the number of sensors is large, sensor data can conflict, and/or costs are associated with sensing. Query generation may require substantial domain knowledge, particularly when different queries can interact. (Destructive testing provides an extreme example) At times, the set of possible queries is large and poorly structured, requiring that complex queries be synthesized rather than selected from a preexisting collection. Finally, expert systems are now being integrated into larger software systems. Such systems may need to access information in other components such as a large data base.

These queries are likely to require substantial computational resources, and should be chosen in an appropriate manner. In typical systems today, queries are chosen in either a depth first or a breadth first manner (relative to the deduction tree). Intelligent query generation will become increasingly more important as expert systems techniques are applied to increasingly complex problems.

Explanation facilities

Explanations are required by both system users and system designers. Users require a rationale for system actions in order to evaluate the utility of those actions. System designers need to be able to monitor and modify the reasoning processes used by the system. The simplest explanation facilities provide an exact trace of the line of reasoning used to come to a decision. (For example, in rule-based systems, an explanation may consist of an ordered list of rule firings) More sophisticated explanations are possible if the line of reasoning is reformulated by mapping it into a level of abstraction more easily comprehended. Major work still needs to be done in accounting for the different audiences requiring explanation capabilities. The heuristics needed to solve complicated problems often make the workings of an expert system difficult to understand, particularly by users not familiar with the internal structure of the system. A tradeoff exists between efficient and comprehensible inference. Comprehensibility can be improved by including in expert systems knowledge structures and inference mechanisms specific to explanation generation.

Joint Industry/University Projects

The researchers are involved in a number of projects which combine the research interests of the SOLVER project and the application interests of companies. These projects include knowledge acquisition and design in the following areas:

Computer system and hardware diagnosis Chemical reactor design Optimal experimental design for new product development and off line quality control.

This kind of joint research is promoted by the Microelectronics and Information Sciences Center of the University of Minnesota (MEIS).

Long range --

Our long range objective is to improve the methodology of the "knowledge capturing" process that occurs in the early stages of the development of expert systems when problem decomposition and solution strategies are being specified. Several related questions of interest include:

What are the performance consequences of different approaches, how can these consequences be evaluated, and what tools can assist in making the best choice?

How can organizations be determined which not only perform well, but are structured so as to facilitate knowledge acquisition from human experts?

We will continue to explore these questions in areas of design and management as well as in law, management, and medicine.

B. Justification and Requirements for Continued SUMEX Use

Our current model development takes advantage of the sophisticated Lisp programming environments on SUMEX and local facilities. Although much current work with Galen is done using a version running on a local VAX 11/780, we continue to benefit from the interaction with other researchers facilitated by the SUMEX system. We expect to use SUMEX to allow other groups access to the Galen program. We also plan to continue use of the knowledge engineering tools available on SUMEX.

We are working toward a Commonlisp implementation of the Galen system and expect to rely heavily on Commonlisp for future projects.

One of our students implemented a demonstration legal expert system in EMYCIN using the SUMEX resource, and we still find that the resource is valuable for making available major systems which we do not have locally, such as EMYCIN.

C. Needs and Plans for Other Computing Resources Beyond SUMEX-AIM

Our current grant from MEIS has permitted us to purchase four Perq 2 AI workstations for our Artificial Intelligence laboratory. The availability of Commonlisp on these machines is one reason why we expect to make use of that language in the future.

SUMEX will continue to be used for collaborative activities and for program development requiring tools not available locally.

D. Recommendations for Future Community and Resource Development

As a remote site, we particularly appreciate the communications that the SUMEX facility provides our researchers with other members of the community. We, too, are moving toward a workstation based development environment, but we hope that SUMEX will continue to serve as a focal point for the medical AI community. In addition to communication and sharing of programs, we are interested in development of Commonlisp based knowledge engineering tools. The continued existence of the SUMEX resource is very important to us.

IV.C. Pilot Stanford Projects

Following are descriptions of the informal pilot projects currently using the Stanford portion of the SUMEX-AIM resource, pending funding, full review, and authorization.

In addition to the progress reports presented here, abstracts for each project are submitted on a separate Scientific Subproject Form.

IV.C.1. REFEREE Project

REFEREE Project

Bruce G. Buchanan, Ph.D., Principal Investigator
Computer Science Department
Stanford University

Byron W. Brown, Ph.D., Co-Principal Investigator
Department of Medicine
Stanford University

Daniel E. Feldman, Ph.D., M.D., Associate Investigator
Department of Medicine
Stanford University

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

The goals of this project are related both to medical science and artificial intelligence: (a) use AI methods to allow the informed but non-expert reader of the medical literature to evaluate a randomized clinical trial, and (b) use the interpretation of the medical literature as a test problem for studies of knowledge acquisition and fusion of information from disparate sources. REFEREE will be used to evaluate the medical literature of clinical trials to determine the quality of a clinical trial, make judgements on the efficacy of the treatment proposed, and synthesize rules of clinical practice. The research is an initial step toward a more general goal - building computer systems to help the clinician and medical scientist read the medical literature more critically and more rapidly.

B. Medical Relevance

The explosive growth of the medical literature has created a severe information gap for the busy clinician. Most physicians can afford neither the time required to study all the pertinent journal articles in their field, nor the risk of ignoring potentially significant discoveries. The majority of clinicians, in fact, have little sophistication in epidemiology and statistics; they must nonetheless base their pragmatic decisions on a combination of clinical experience and published literature. The clinician's computerized assistant must ferret out useful maxims of clinical practice from the medical literature, pass judgment on the quality of medical reports, evaluate the efficacy of proposed treatments, and adjudicate the interpretation of conflicting and even contradictory studies.

C. Highlights of Progress

REFEREE, a rule-based system built upon the EMYCIN framework, partially encodes the epidemiological knowledge of two highly regarded experts at Stanford, a biostatistician (Dr. Bill Brown) and a clinician (Dr. Dan Feldman). The REFEREE system, in particular, allows the informed but non-expert reader of the medical literature to study the quality, believability, statistical precision, and overall merit of a randomized clinical trial. REFEREE structures its analysis into five categories: the CREDENTIALS of the author and the institution behind the study; the PLANNING of

the trial, including stopping rules, selection of subgroups, and standards for evaluating the endpoints; the EXECUTION of the study, including the degree of blinding and the adherence to planning standards; the REPORTING of results; and the statistical ANALYSIS applied to those results.

In the future, REFEREE will alleviate the knowledge-acquisition bottleneck for an automated medical decision-maker: the program will evaluate the quality of a clinical trial, judge the efficacy of the treatment proposed therein, and synthesize rules of clinical practice. For the present, however, the fusion of knowledge from disparate sources remains a problem in pure AI. The system's designers, Dr. Bruce Buchanan and R. Martin Chavez, have instead focused their efforts on the refinement and deepening of REFEREE's epidemiological knowledge.

In its present incarnation, REFEREE holds its knowledge of controlled clinical methodology in a database of 250 rules. The investigators working on the REFEREE project have, over the last year, produced a prototype consultant that evaluates the design and reporting of a randomized control trial. Clearly, REFEREE must possess a great deal of statistical sophistication before it can usefully serve the critical reader of medical literature.

In the spring of 1986, Chavez undertook the task of structuring the vast corpus of statistical techniques used in clinical trials. The current version of REFEREE contains, in preliminary form, Prof. Brown's expert knowledge of biostatistics. REFEREE evaluates each statistical procedure described by the authors of the paper. The automated consultant then determines the most appropriate method for the problem at hand, based on the design of the trial and the hypotheses to be tested. REFEREE checks critical assumptions, looks for possible statistical abuses, verifies adjustments, and re-computes the statistics. In a beta-blocker study that employs the Cox proportional-hazards model, for instance, REFEREE will analyze the Kaplan-Meier survival curve and verify or reject the presence of a significant treatment effect.

In order to evaluate the paper's presentation of a statistical test, REFEREE must apply three kinds of knowledge. First, REFEREE examines the problem at hand and exhaustively establishes a list of relevant techniques. REFEREE applies rules of the following form in order to build the list of appropriate tests:

> IF: 1) the data consist of two independent random samples, and
> 2) A: outlying data prevent the application of a test based
>   on sample means, or
>   B: it is not known whether the data are from normally-
>   distributed populations,
>   and
> 3) no censored data are present,
> THEN: The nonparametric ranktest of Wilcoxon is a reasonable test.

REFEREE presently possesses about fifty rules that define the range of admissible tests.

Second, REFEREE must know how to perform each kind of statistical test in order to verify the paper's results. Algorithmic knowledge (expressed as Interlisp code), for instance, enables REFEREE to calculate chi-squared for JxK contingency tables, evaluate the binomial distribution, and compute the Gehan rank statistic for survival curves. REFEREE knows how to apply the majority of statistical methods in common usage.

Finally, REFEREE must know how to test hypotheses. In order to look for significance in a contingency table, for instance, REFEREE must know how to compute P values and critical values from the chi-squared distributions. REFEREE can understand and interpret all the major distributions used in clinical epidemiology.

REFEREE may therefore determine on its own whether the authors' data support their conclusions.

The interaction between general techniques and specific clinical trials poses the most significant obstacle to REFEREE's success. The automated consultant must capture the ill-defined knowledge that enables the expert biostatistician to reason from theoretical principles to specific applications and interpretations in the field of medicine. In addition, REFEREE must produce, in clear English prose, a comprehensive critique of every clinical trial submitted for analysis. Future versions of REFEREE must synthesize those analyses into rules for an automated medical decision maker.

At present, REFEREE enables the clinician to read clinical trials more critically. As such, REFEREE represents only the first step in a larger research plan, the automation of knowledge acquisition. Current work in the restricted domain of clinical trials will, we hope, illustrate general principles in the design of decision makers that gather expertise from written text and multiple knowledge sources.

D. Relevant Publications Haggerty, J.: *REFEREE and RULECRITIC: Two prototypes for assessing the quality of a medical paper.* REPORT KSL-84-49. Master's Thesis, Stanford University, May 1984.

E. Funding Support REFEREE currently has no funding. The personnel involved donate their time to this project.

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaborations

Dr. Brown and Dr. Feldman of the Stanford University School of Medicine are actively involved in the REFEREE project and are the primary domain experts for this project.

B. Interactions with other SUMEX-AIM projects

C. Critique of Resource Management The SUMEX computer resource has been essential for the work to date, and the SUMEX staff has continued to be very cooperative with the REFEREE project.

III. RESEARCH PLANS

A. Goals & Plans REFEREE will consist of three separate but closely related programs:

1. System I will assist in the evaluation of the *quality* of a *single* clinical trial. The user will be imagined to be the editor of a journal reviewing a manuscript for publication, but the program will be tested on a variety of readers, including clinicians, medical scientists, medical and graduate students, and clerical help.

2. System II will assist in the evaluation of the *effectiveness* of the treatment or intervention examined in a *single* published clinical trial. The user will be imagined to be a clinician interested in judging the efficacy of the treatment being tested in the trial.

3. System III will assist in the evaluation of the *effectiveness* of a single treatment examined in a *number* of published clinical trials.

We also intend to test the first two systems against unassisted evaluations by the various categories of readers. The testing will include a formal testing of the programs by comparing the speed and number of flaws found in using the program with similar

measurements on unassisted reading. In addition there will be a more informal evaluation by questionnaire of the subjective impressions of users of the program, ascertaining the likelihood of routine use and the value of such a program to the user.

B. Justification for continued SUMEX use The DEC 2060 of the SUMEX computing resource is currently the only system we have available for developing REFEREE in Interlisp and the EMYCIN framework.

C. Need for other computing resources The REFEREE project needs a high performance LISP workstation to assist in the development and execution of the REFEREE programs. We believe that such a machine is important in order to explore the issues of making the programs easy to use through graphics and a good user interface, as well as to make the programs easily exportable.

IV.D. Pilot AIM Projects

Following is a description of the informal pilot project currently using the AIM portion of the SUMEX-AIM resource, pending funding, full review, and authorization.

In addition to the progress report presented here, an abstract is submitted on a separate Scientific Subproject Form.

IV.D.1. PATHFINDER Project

PATHFINDER Project

Bharat Nathwani, M.D.
Department of Pathology
University of Southern California

Lawrence M. Fagan, M.D., Ph.D.
Department of Medicine
Stanford University

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

Our project addresses difficulties in the diagnosis of lymph node pathology. Five studies from cooperative oncology groups have documented that, while experts show agreement with one another, the diagnosis made by practicing pathologists may have to be changed by expert hematopathologists in as many as 50% of the cases. Precise diagnoses are crucial for the determination of optimal treatment. To make the knowledge and diagnostic reasoning capabilities of experts available to the practicing pathologist, we have developed a pilot computer-based diagnostic program called PATHFINDER. The project is a collaborative effort of the University of Southern California and the Stanford University Medical Computer Science Group. A pilot version of the program provides diagnostic advice on 72 common benign and malignant diseases of the lymph node based on 110 histologic features. Our research plans are to develop a full-scale version of the computer program by substantially increasing the quantity and quality of knowledge and to develop techniques for knowledge representation and manipulation appropriate to this application area. The design of the program has been strongly influenced by the INTERNIST/CADUCEUS program developed on the SUMEX resource.

A group of expert pathologists from several centers in the U.S. have showed interest in the program and helped to provide the structure of the knowledge base for the PATHFINDER system.

B. Medical Relevance and Collaboration

One of the most difficult areas in surgical pathology is the microscopic interpretation of lymph node biopsies. Most pathologists have difficulty in accurately classifying lymphomas. Several cooperative oncology group studies have documented that while experts show agreement with one another, the diagnosis rendered by a "local" pathologist may have to be changed by expert lymph node pathologists (expert hematopathologists) in as many as 50% of the cases.

The National Cancer Institute recognized this problem in 1968 and created the Lymphoma Task Force which is now identified as the Repository Center and the Pathology Panel for Lymphoma Clinical Studies. The main function of this expert panel of pathologists is to confirm the diagnosis of the "local" pathologists and to ensure that the pathologic diagnosis is made uniform from one center to another so that the comparative results of clinical therapeutic trials on lymphoma patients are valid. An expert panel approach is only a partial answer to this problem. The panel is useful in only a small percentage (3%) of cases; the Pathology Panel annually reviews

only 1,000 cases whereas more than 30,000 new cases of lymphomas are reported each year. A Panel approach to diagnosis is not practical and lymph node pathology cannot be routinely practiced in this manner.

We believe that practicing pathologists do not see enough case material to maintain a high-level of diagnostic accuracy. The disparity between the experience of expert hematopathology teams and those in community hospitals is striking. An experienced hematopathology team may review thousands of cases per year. In contrast, in a community hospital, an average of only ten new cases of malignant lymphomas are diagnosed each year. Even in a university hospital, only approximately 100 new patients are diagnosed every year.

Because of the limited numbers of cases seen, pathologists may not be conversant with the differential diagnoses consistent with each of the histologic features of the lymph node; they may lack familiarity with the complete spectrum of the histologic findings associated with a wide range of diseases. In addition, pathologists may be unable to fully comprehend the conflicting concepts and terminology of the different classifications of non-Hodgkin's lymphomas, and may not be cognizant of the significance of the immunologic, cell kinetic, cytogenetic, and immunogenetic data associated with each of the subtypes of the non-Hodgkin's lymphomas.

In order to promote the accuracy of the knowledge base development we will have participants for multiple institutions collaborating on the project. Dr. Nathwani will be joined by experts from Stanford (Dr. Dorfman), St. Jude's Children's Research Center -- Memphis (Dr. Berard) and City of Hope (Dr. Burke).

C. Highlights of Research Progress

C.1 Previous Accomplishments

Since the project's inception in September, 1983, we have constructed several versions of PATHFINDER. The first several versions of the program were *rule-based* systems like MYCIN and ONCOCIN which were developed earlier by the Stanford group. We soon discovered, however, that the large number of overlapping features in diseases of the lymph node would make a rule-based system cumbersome to implement. We next considered the construction of a *hybrid system*, consisting of a rule-based algorithm that would pass control to an INTERNIST-like scoring algorithm if it could not confirm the existence of classical sets of features. We finally decided that a modified form of the INTERNIST program would be most appropriate. The original version of PATHFINDER is written in the computer language Maclisp and runs on the SUMEX DEC-20. This was transferred to Portable Standard Lisp (PSL) on the DEC-20, and later transferred to PSL on the HP 9836 workstations. Two graduate students, David Heckerman and Eric Horvitz, designed and implemented the program.

The prototype knowledge base was constructed by Dr. Nathwani. During the early part of 1984, we organized two meetings of the entire team, including the pathology experts, to define the selection of diseases to be included in the system, and the choice of features to be used in the scoring process.

During the last year, we have focused on methodologies for more accurately representing expert beliefs. In particular, we have used *influence diagrams* to represent dependencies among features in the PATHFINDER knowledge base. A great deal of effort has been devoted to assessing and representing the intricate relationships among features that exist in the domain. We believe that this process will help to overcome some of the limitations of medical diagnostic systems.

C.1 The PATHFINDER knowledge base

The basic building block of the PATHFINDER knowledge base is the disease profile or

frame. Each disease frame consists of *features* useful for diagnosis of lymph node diseases. Currently these features include histopathologic findings seen in both low- and high-power magnifications. Each feature is associated with a list of exhaustive and mutually exclusive *values*. For example, the feature *pseudofollicularity* can take on any one of the values *absent, slight, moderate,* or *prominent.* These lists of values give the program access to *severity* information. In addition, these lists eliminate obvious interdependencies among the values for a given feature. For example, if pseudofollicularity is *moderate,* it cannot also be *absent.*

Qualitative dependencies among features for each disease are represented using the influence diagram methodology mentioned above. An influence diagram contains *nodes* and *arcs.* Nodes represent features and arcs represent dependencies among features. In particular, an arc is drawn from one feature to another when an expert believes that knowing one feature can change his beliefs that another feature will take on its possible values even when the diagnosis is known. Probabilities are used to quantitate the beliefs asserted by the expert.

C.2 New Hewlett Packard Workstation

Through the USC-affiliated Information Sciences Institute, Dr. Nathwani has obtained a Hewlett-Packard Workstation that is similar to the 9836. The Pathfinder program has been brought up on this machine. This means that the program now exists on three different machines, in three separate locations, using one standard language (Portable Standard Lisp). Thus, the need for support of networked machines and communications has increased during this last year.

D. Publications Since January 1984

Horvitz, E.J., Heckerman, D.E., Nathwani, B.N. and Fagan, L.M.: *Diagnostic Strategies in the Hypothesis-directed PATHFINDER System, Node Pathology.* HPP Memo 84-13. Proceedings of the First Conference on Artificial Intelligence Applications, Denver, Colorado, Dec., 1984.

Horvitz, E.J., Heckerman, D.E., Nathwani, B.N. and Fagan, L.M.: *The Use of a Heuristic Problem-solving Hierarchy to Facilitate the Explanation of Hypothesis-directed Reasoning.* KSL Memo 86-2. Proceedings of MedInfo, Washington D.C., October, 1986.

E. Funding Support

Research Grant submitted to National Institutes of Health
Grant Title: "Computer-aided Diagnosis of Malignant Lymph Node Diseases"
Principal Investigator: Bharat Nathwani
Funding pending from the National Library of Medicine

Professional Staff Association, Los Angeles County Hospital, $10,000.

University of Southern California, Comprehensive Cancer Center, $30,000.

Project Socrates, Univ. of Southern Calif., Gift from IBM of IBM PC/XT.

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaborations and Program Dissemination via SUMEX

Because our team of experts are in different parts of the country and the computer

scientists are not located at the USC, we envision a tremendous use of SUMEX for communication, demonstration of programs, and remote modification of the knowledge base. The proposal mentioned above was developed using the communication facilities of SUMEX.

B. *Sharing and Interaction with Other SUMEX-AIM Projects*

Our project depends heavily on the techniques developed by the INTERNIST/CADUCEUS project. We have been in electronic contact and have met with members of the INTERNIST/CADUCEUS project, as well as been able to utilize information and experience with the INTERNIST program gathered over the years through the AIM conferences and on-line interaction. Our experience with the extensive development of the pathology knowledge base utilizing multiple experts should provide for intense and helpful discussions between our two projects.

The SUMEX pilot project, RXDX, designed to assist in the diagnosis of psychiatric disorders, is currently using a version of the PATHFINDER program on the DEC-20 for the development of early prototypes of future systems.

C. *Critique of Resource Management*

The SUMEX resource has provided an excellent basis for the development of a pilot project. The availability of a pre-existing facility with appropriate computer languages, communication facilities (especially the TYMNET network), and document preparation facilities allowed us to make good progress in a short period of time. The management has been very useful in assisting with our needs during the start of this project.

III. RESEARCH PLANS

A. *Project Goals and Plans*

Collection and refinement of knowledge about lymph node pathology

The knowledge base of the program is about to undergo revision by the experts, and then will be extensively tested. A logical next step would be to extend the program to clinical settings, as well as possible extensions of the knowledge base.

Other possible extensions include: developing techniques for simplifying the acquisition and verification of knowledge from experts, creating mapping schemes that will facilitate the understanding of the many classifications of non-Hodgkin's lymphomas. We will also attempt to represent knowledge about special diagnostic entities, such as multiple discordant histologies and atypical proliferations, which do not fit into the classification methods we have utilized.

Representation Research

We hope to enhance the INTERNIST-1 model by structuring features so that overlapping features are not incorrectly weighted in the decision making process, implementing new methods for scoring hypotheses, and creating appropriate explanation capabilities.

B. *Requirements for Continued SUMEX Use*

We are currently dependent on the SUMEX computer for the use of the program by remote users, and for project coordination. We have transferred the program over to Portable Standard Lisp which is used by several users on the SUMEX system. While the switch to workstations has lessened our requirements for computer time for the development of the algorithms, we will continue to need the SUMEX facility for the

interaction with each of the research locations specified in our NIH proposal. The HP equipment is currently unable to allow remote access, and thus the program will have to be maintained on the 2060 for use by all non-Stanford users.

C. Requirements for Additional Computing Resources

Most of our computing resources will be met by the 2060 plus the use of the HP9836 workstation. We will need additional file space on the 2060 as we quadruple the size of our knowledge base. We will continue to require access to the 2060 for communication purposes, access to other programs, and for file storage and archiving.

D. Recommendations for Future Community and Resource Development

We encourage the continued exploration by SUMEX of the interconnection of workstations within the mainframe computer setting. We will need to be able to quickly move a program from workstation to workstation, or from workstation back and forth to the mainframe. Software tools that would help the transfer of programs from one type of workstation to another would also be quite useful. Until the type of workstations that we are using in this research becomes inexpensive ($5000 or less), we will continue to need a machine like SUMEX to provide others with a chance to experiment with our software.

E. H. Shortliffe

IV.D.2. RXDX Project

RXDX Project

Robert Lindsay, Ph.D.
Michael Feinberg, M.D., Ph.D.
Manfred Kochen, Ph.D.
University of Michigan
Ann Arbor, Michigan

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

We are developing a prototype expert system that could act as a consultant in the diagnosis and management of depression. Health professionals will interact with the program as they might with a human consultant, describing the patient, receiving advice, and asking the consultant about the rationale for each recommendation. The program uses a knowledge base constructed by encoding the clinical expertise of a skilled psychiatrist in a set of rules and other knowledge structures. It will use this knowledge base to decide on the most likely diagnosis (endogenous or nonendogenous depression), assess the need for hospitalization, and recommend specific somatic treatments when this is indicated (e.g., tricyclic antidepressants). The treatment recommendation will take into account the patient's diagnosis, age, concurrent illnesses, and concurrent treatments (drug interactions).

B. Medical Relevance and Collaboration

There is a documented shortage of psychiatrists in the US (GMENAC, 1980), and the estimates of the prevalence of psychiatric illness used to develop that report were lower than the figures in recent population surveys (Myers et al., 1984). Further, most prescriptions for antidepressants are written by non-psychiatrists (Johnson, 1974; Kline, 1974) and the great majority of depressed patients seen by a sample of primary care physicians were treated inappropriately (Weissman et al., 1981). These data highlight the need for improving the treatment provided to the majority of mentally ill patients. We believe that computers can act as consultants to non-psychiatrist clinicians, resulting in improved patient care.

The potential benefits to Psychiatry include: making relatively skilled psychiatric consultation widely available in underserved areas, including some public mental health facilities where patients are seen by non-psychiatrists and have relatively little direct patient-physician contact; providing non-psychiatrically trained physicians with additional information about psychiatric diagnosis and treatment; avoiding errors of oversight caused by inaccessible patient data; and increased productivity in patient care. Like any good consultant, the program will be able to teach the interested user, and can function as a teaching tool independent of direct clinical application.

We examined two other SUMEX-based psychiatry projects, the BLUEBOX project of Mulsant and Servan-Schreiber (1984), and the HEADMED project of Heiser and Brooks (1978, 1980). Mulsant and Servan-Schreiber visited us at Michigan and discussed the rationale and progress of their project. Heiser also visited with us and agreed to collaborate with our project as a consultant. This relationship ended with the termination of our NIMH Small Grant in Sept. 1985. We have discussed possible

collaboration with William Swartout, who plans to expand his explanation system methodology into other medical domains as part of the process of making the method more general.

C. Highlights of Research Progress

During the past year, we have concentrated on two separate expert systems. The first is an expert on diagnosis, specifically in distinguishing endogenous (i.e., drug-responsive) depression from nonendogenous depression. We are also adding the knowledge needed to separate other subtypes of depression. We are using the Pathfinder program written by Fagan and his co-workers (Horvitz et al., 1984) with their gracious permission. The knowledge base for this program contains data on the prevalence of each illness in the population of interest and the frequency of occurrence of each symptom in each illness. These data are used in a modification of Bayes' rule to reach a diagnosis. The program performs very well (Feinberg and Lindsay, in press). We have been working on SUMEX and are now trying to move this program to a personal computer to make it more widely available.

Our major project during the past year has been an expert system for the somatic treatment of (endogenous) depression, where somatic treatment includes antidepressant drugs, electroshock, and lithium. We are writing this system using KEE, an expert system shell generously donated by Intellicorp, running on a Xerox 1108 workstation. We have been able to incorporate the work we did earlier on SUMEX, either directly by transporting the rules or indirectly by using what we learned about building expert systems in general. The knowledge base includes information about the side effects of each of the drugs and about the physiological mechanisms of these side effects. This information allows us to predict drug interactions and the likelihood of occurrence of various side effects in a given patient, and to base explanations on knowledge of the underlying physiology. The knowledge base also includes specific information about drug regimens, about preventing and treating side effects, and about how to take all of this into account in selecting a drug and dosage regimen for the individual patient.

D. List of Relevant Publications

1. We have one publication in press: (*) Feinberg, M. and Lindsay, R. K.: *Expert systems*. Psychopharmacol. Bull., in press.

2. We also presented our work at the World Congress of Biological Psychiatry, Philadelphia, Sept. 1985. : (*) Feinberg, M. and Lindsay, R. K.: *Expert systems*.

3. We have submitted an abstract to AAAI 1986 and to MedInfo and are awaiting a response. We will be presenting our work in a panel at MedInfo.

E. Funding Support

We have received support from the Vice-President for Research at the University of Michigan, and from the NIH "Small Grants" Program (Grant Number R03MH40239-01; Total Direct Costs = $13,850). These funds have enabled us to gather the pilot data for grant applications submitted to NIH.

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaboration and Program Dissemination via SUMEX

We have established via SUMEX a community of researchers who are interested in AI applications in psychiatry. We also have used the message system to communicate with other AI scientists at SUMEX and elsewhere. As noted above, we have transferred the Pathfinder program from HP workstations to an IBM Model AT via SUMEX.

B. Sharing and Collaboration with other SUMEX-AIM Projects

Our use of EMYCIN and AGE has been of major importance. In addition, we have worked with Dr. Larry Fagan to learn about his Pathfinder program. We used that program, on SUMEX, to obtain some information for the RxDx project by applying it to data we previously collected on depression symptom frequencies. As described above, we are continuing to use the Pathfinder program, both on SUMEX and on a microcomputer.

C. Critique of Resource Management

We have been using EMYCIN, AGE, and Pathfinder in our work, and have found these programs very valuable, saving us many hours of programming in LISP. There are some problems with the first two, many of which center around discrepancies between the versions described in the manuals and the versions actually running on SUMEX. We would suggest that software be more strongly supported than is now the case, so that it and SUMEX can be even more useful to beginners in AI in Medicine.

SUMEX itself has been invaluable. We don't have ready access to any other machine of equal computing power which also has a strongly supported LISP available. Specifically, the LISP compiler available on the Amdahl 5860 here differs from those used at major AI centers such as Stanford and MIT. We have also made good use of the ARPANET connections that SUMEX offers. Feinberg spent a month of his sabbatical working with Prof. Peter Szolovits at MIT, learning about AI in Medicine. This visit was arranged using computer mail through SUMEX. Lindsay and Feinberg were able to continue their collaborative work while the latter was in Cambridge, using the same medium. The alternative would have been days lost in the mails and many dollars spent on phone calls.

Most of the limitations of SUMEX, and they are often severe, derive from the necessity to access it via TYMNET. Response time is often impossibly slow, and even at its best the delays are annoying and frustrating, even for editing and debugging. SUMEX has just been made available through UNINET, and this seems much better after about two weeks of use.

III. RESEARCH PLAN

A. Project Goals and Plans

Our immediate objective is to develop expert systems that can differentiate patients with the various subtypes of depressive disorder, and prescribe appropriate treatment. This system should perform at about the level of a board-certified psychiatrist, i.e. better than an average resident but not as well as a human expert in depression. Eventually, we plan to enlarge the knowledge base so that the expert system can diagnose and prescribe for a wider range of psychiatric patients, particularly those with illnesses that are likely to respond to psychopharmacological agents. We will design the system so that it could be used by non-medical clinicians or by non-psychiatrist MD's as an adjunct to consultation with a human expert. We plan also to focus on problems of the user interface and the integration of this system with other databases.

B. Justification and Requirements for continued SUMEX use

The access to SUMEX resources is essentially our sole means of maintaining contact with the community of researchers working on applications of AI in medicine. Although we are moving our system to local workstations, the communications capability of SUMEX will continue to be important.

We anticipate that our requirements for computing time and file space will continue at about the same level for the next year.

C. Needs and Plans for Other Computing Resources

As our project evolves and we run into the limitations of the time-shared SUMEX facility, we anticipate employing different expert systems software. We are now using KEE and Golden Common LISP, the former on a Xerox 1108; the latter on an IBM Model AT. Ultimately, when our consultant is made available for field trials and clinical use, it will need to be transported to a personal computer that is large enough to support the system yet inexpensive enough to be widely available. Both KEE and common LISP will now run on microcomputers, the former requiring a more powerful central machine as well. The cost of the hardware needed to run our systems has dropped by half or more in one year, and is continuing to drop. Our design strategy is to avoid limiting ourselves and our aspirations to that which is affordable today; instead we will attempt to project the growth of our project and the price-performance curve of computing such that they meet at some reasonable point in the future.

D. Recommendations for Future Community and Resource Development

Valuable as the present SUMEX facilities are to us, they are in many ways limited and awkward to use. The major limitation we feel is the difficulty and sometimes the impossibility of making contact with everyone who could be of value to us. We hope that greater emphasis will be put on internetwork gateways. It is important not only to establish more of these, but to develop consistent and convenient standards for electronic mail, electronic file transfers, graphic information transfer, national archives and data bases, and personal filing and retrieval (categorization) systems. The present state of the art feels quite limiting, now that the basic concepts of computer networking have become available and have proved their potential.

We expect that the role of the SUMEX-AIM resource will continue to evolve in the direction of increased importance of communication, including graphical information, electronic dissemination of preprints, and database and program access. The need for computer cycles on a large mainframe will diminish. We hope to have continued access to the system for communication, but do not anticipate continued use of it as a LISP computation server beyond the next year or eighteen months.

If fees for using SUMEX resources were imposed, this would have a drastically limiting effect on the value of the system to us. Even if we had a budget to purchase such services, the inhibiting effect of having a meter running would cause us to make less use of it that we should. We have been conscious of the costs of the system and feel that we have not used it imprudently, even though we have not directly borne its costs.

E. H. Shortliffe

Appendix A

AIM Management Committee Membership

Following are the current membership lists of the various SUMEX-AIM management committees:

AIM Executive Committee:

SHORTLIFFE, Edward H., M.D., Ph.D.   (Chairman)
       Principal Investigator - SUMEX
       Medical Computer Science, TC135
       Stanford University Medical Center
       Stanford, California 94305
       (415) 497-6970

FEIGENBAUM, Edward A., Ph.D.
       Co-Principal Investigator - SUMEX
       Heuristic Programming Project
       Department of Computer Science
       701 Welch Road, Building C
       Stanford University
       Stanford, California 94305
       (415) 497-4879

KULIKOWSKI, Casimir, Ph.D.
       Department of Computer Science
       Rutgers University
       New Brunswick, New Jersey 08903
       (201) 932-2006

LEDERBERG, Joshua, Ph.D.
       President
       The Rockefeller University
       1230 York Avenue
       New York, New York 10021
       (212) 570-8080, 570-8000

LINDBERG, Donald A.B., M.D.   (Past Adv Grp Chrmn)
       Director, National Library of Medicine
       8600 Rockville Pike
       Bethesda, Maryland 02114
       (617) 726-8311

MYERS, Jack D., M.D.
       School of Medicine
       Scaife Hall, 1291
       University of Pittsburgh
       Pittsburgh, Pennsylvania 15261
       (412) 624-2649

E. H. Shortliffe

AIM Advisory Group:

MYERS, Jack D., M.D.   (Chairman)
         School of Medicine
         Scaife Hall, 1291
         University of Pittsburgh
         Pittsburgh, Pennsylvania 15261
         (412) 624-2649

AMAREL, Saul, Ph.D.
         Department of Computer Science
         Rutgers University
         New Brunswick, New Jersey 08903
         (201) 932-3546

COULTER, Charles L., Ph.D.   (Exec. Secretary)
         Bldg 31, Room 5B41
         Biomedical Research Technology Program
         National Institutes of Health
         Bethesda, Maryland 20205

FEIGENBAUM, Edward A., Ph.D.   (Ex-officio)
         Co-Principal Investigator - SUMEX
         Heuristic Programming Project
         Department of Computer Science
         701 Welch Road, Building C
         Stanford University
         Palo Alto, California 94305
         (415) 497-4879

KULIKOWSKI, Casimir, Ph.D.
         Department of Computer Science
         Rutgers University
         New Brunswick, New Jersey 08903
         (201) 932-2006

LEDERBERG, Joshua, Ph.D.
         President
         The Rockefeller University
         1230 York Avenue
         New York, New York 10021
         (212) 570-8080, 570-8000

LINDBERG, Donald A.B., M.D.
         Director, National Library of Medicine
         8600 Rockville Pike
         Bethesda, Maryland 02114
         (617) 726-8311

MINSKY, Marvin, Ph.D.
         Artificial Intelligence Laboratory
         Massachusetts Institute of Technology
         545 Technology Square
         Cambridge, Massachusetts 02139
         (617) 253-5864

MOHLER, William C., M.D.
　　　　Associate Director
　　　　Division of Computer Research and Technology
　　　　National Institutes of Health
　　　　Building 12A, Room 3033
　　　　9000 Rockville Pike
　　　　Bethesda, Maryland 20205
　　　　(301) 496-1168

PAUKER, Stephen G., M.D.
　　　　Department of Medicine - Cardiology
　　　　Tufts New England Medical Center Hospital
　　　　171 Harrison Avenue
　　　　Boston, Massachusetts 02111
　　　　(617) 956-5910

SHORTLIFFE, Edward H., M.D., Ph.D.　　　(Ex-officio)
　　　　Principal Investigator - SUMEX
　　　　Medical Computer Science, TC135
　　　　Stanford University Medical Center
　　　　Stanford, California 94305
　　　　(415) 497-6970

SIMON, Herbert A., Ph.D.
　　　　Department of Psychology
　　　　Baker Hall, 339
　　　　Carnegie-Mellon University
　　　　Schenley Park
　　　　Pittsburgh, Pennsylvania 15213
　　　　(412) 578-2787, 578-2000

Stanford Community Advisory Committee:

FEIGENBAUM, Edward A., Ph.D.   (Chairman)
   Heuristic Programming Project
   Department of Computer Science
   Margaret Jacks Hall
   Stanford University
   Stanford, California 94305
   (415) 497-4879

LEVINTHAL, Elliott C., Ph.D.
   Departments of Mechanical and Electrical Engineering
   Building 530
   Stanford University
   Stanford, California 94305
   (415) 497-9037

SHORTLIFFE, Edward H., M.D., Ph.D.
   Principal Investigator - SUMEX
   Medical Computer Science, TC135
   Stanford University Medical Center
   Stanford, California 94305
   (415) 497-6970

Appendix B
Scientific Subproject Abstracts

The following are brief abstracts of our collaborative research projects.

Stanford Project: GUIDON/NEOMYCIN --
                                 KNOWLEDGE ENGINEERING
                                 FOR TEACHING MEDICAL DIAGNOSIS

Principal Investigators: William J. Clancey, Ph.D.
                                 701 Welch Road
                                 Department of Computer Science
                                 Stanford University
                                 Palo Alto, California 94304
                                 (415) 497-1997 (CLANCEY@SUMEX-AIM)

                                 Bruce G. Buchanan, Ph.D.
                                 Computer Science Department
                                 70l Welch Road
                                 Stanford University
                                 Palo Alto, California 94304
                                 (415) 497-0935 (BUCHANAN@SUMEX-AIM)

SOFTWARE AVAILABLE ON SUMEX

GUIDON--A system developed for intelligent computer-aided instruction. Although it was developed in the context of MYCIN's infectious disease knowledge base, the tutorial rules will operate upon any EMYCIN knowledge base.

NEOMYCIN--A consulation system derived from MYCIN, with the knowledge base greatly extended and reconfigured for use in teaching. In contrast with MYCIN, diagnostic procedures, common sense facts, and disease hierarchies are factored out of the basic finding/disease associations. The diagnostic procedures are abstract (not specific to any problem domain) and model human reasoning, unlike the exhaustive, top-down approach implicit in MYCIN's medical rules. This knowledge base will be used in the GUIDON2 family of instructional programs, being developed on D-machines.

REFERENCES

Clancey, W.J.: *Overview of GUIDON*. In A. Barr and E.A. Feigenbaum (Eds.), THE HANDBOOK OF ARTIFICIAL INTELLIGENCE, Vol. 2. William Kaufmann Assoc., Los Altos, CA, 1982. (Also to appear in Journal of Computer-based Instruction)

Clancey, W.J.: *Methodology for building an intelligent tutoring system.* In Kintsch, Polson, and Miller, (Eds.), METHODS AND TACTICS IN COGNITIVE SCIENCE. L. Erlbaum Assoc., Hillsdale, NJ. 1984. (Also STAN-CS-81-894, HPP 81-18)

Clancey, W.J.: *Acquiring, representing, and evaluating a competence model of diagnosis*. In Chi, Glaser, and Farr (Eds.), THE NATURE OF EXPERTISE. In preparation. HPP-84-2.

Stanford Project: MOLGEN -- AN EXPERIMENT PLANNING SYSTEM
                                 FOR MOLECULAR GENETICS

Principal Investigators: Edward A. Feigenbaum, Ph.D.
                                 Department of Computer Science
                                 Stanford University

                                 Charles Yanofsky, Ph.D. (YANOFSKY@SUMEX-AIM)
                                 Department of Biology
                                 Stanford University
                                 Stanford, California 94305
                                 (415) 497-2413

                                 Contact: Dr. Peter FRIEDLAND@SUMEX-AIM
                                 (415) 497-3728

The goal of the MOLGEN Project is to apply the techniques of artificial intelligence to the domain of molecular biology with the aim of providing assistance to the experimental scientist. Previous work has focused on the task of experiment design. Two major approaches to this problem have been explored, one which instantiates abstracted experimental strategies with specific laboratory tools, and one which creates plans in toto, heavily influenced by the role played by interactions between plan steps. As part of the effort to build an experiment design system, a knowledge representation and acquisition package--the UNITS System, has been constructed. A large knowledge base, containing information about nucleic acid structures, laboratory techniques, and experiment-design strategies, has been developed using this tool. Smaller systems, such as programs which analyze primary sequence data for homologies and symmetries, have been built when needed.

New work has begun on scientific theory formation, modification, and testing. This work will be done within the domain of regulatory genetics. We plan to explore fundamental issues in machine learning and discovery, as well as construct systems that will assist the laboratory scientist in accomplishing his intellectual goals.

SOFTWARE AVAILABLE ON SUMEX

    SPEX system for experiment design.
    UNITS system for knowledge representation and acquisition.
    SEQ system for nucleotide sequence analysis.

REFERENCES

Friedland, P.E.: *Knowledge-based experiment design in molecular genetics*, (Ph.D. thesis). Stanford Computer Science Report, STAN-CS-79-771.

Friedland, P.E. and Iwasaki, Y.: *The concept and implementation of skeletal plans*, Journal of Automated Reasoning, 1(2):161-208, 1985.

Friedland, P.E. and Kedes, L.: *Discovering the secrets of DNA*, Communications of the ACM, 28(11):1164-1186, November, 1985.

Stefik, M.J.: *An examination of a frame-structured representation system*, Proc. Sixth IJCAI, Tokyo, August, 1979, pp. 845-852.

Stefik, M.J.: *Planning with constraints*, (Ph.D. thesis). Stanford Computer Science Report, STAN-CS-80-784, March, 1980.

Stanford Project: ONCOCIN -- KNOWLEDGE ENGINEERING FOR
                             ONCOLOGY CHEMOTHERAPY CONSULTATION

Principal Investigator: Edward H. Shortliffe, M.D., Ph.D.
                             Departments of Medicine and Computer Science
                             Stanford University Medical Center
                             Medical School Office Building
                             Stanford, California 94305
                             (415) 723-6979 (SHORTLIFFE@SUMEX-AIM)

Project Director: Dr. Lawrence M. Fagan

The ONCOCIN Project is overseen by a collaborative group of physicians and computer scientists who are developing an intelligent system that uses the techniques of knowledge engineering to advise oncologists in the management of patients receiving cancer chemotherapy. The general research foci of the group members include knowledge acquisition, inexact reasoning, explanation, and the representation of time and of expert thinking patterns. Much of the work developed from research in the 1970's on the MYCIN and EMYCIN programs, early efforts that helped define the group's research directions for the coming decade. MYCIN and EMYCIN are still available on SUMEX for demonstration purposes.

The prototype ONCOCIN system is in limited experimental use by oncologists in the Stanford Oncology Clinic. Thus much of the emphasis of this research has been on human engineering so that the physicians will accept the program as a useful adjunct to their patient care activities. ONCOCIN has generally been well-accepted since its introduction, and we are now testing a version of the program which runs on professional workstations (rather than the central SUMEX computer) so that it can be implemented and evaluated at sites away from the University.

SOFTWARE AVAILABLE ON SUMEX

MYCIN--   A consultation system designed to assist physicians with the selection of antimicrobial therapy for severe infections. It has achieved expert level performance in formal evaluations of its ability to select therapy for bacteremia and meningitis. Although MYCIN is no longer the subject of an active research program, the system continues to be available on SUMEX for demonstration purposes and as a testing environment for other research projects.

EMYCIN--   The "essential MYCIN" system is a generalization of the MYCIN knowledge representation and control structure. It is designed to facilitate the development of new expert consultation systems for both clinical and non-medical domains.

ONCOCIN--   This system is in clinical use but requires lisp machines to be run. Much of the knowledge in the domain of cancer chemotherapy is already well-specified in protocol documents, but expert judgments also need to be understood and modeled.

REFERENCES

Shortliffe, E.H., Scott, A.C., Bischoff, M.B., Campbell, A.B., van Melle, W. and Jacobs, C.D.: *ONCOCIN: An expert system for oncology protocol management*. Proc. Seventh IJCAI, pp. 876-881, Vancouver, B.C., August, 1981.

Duda, R.O. and Shortliffe, E.H.: *Expert systems research*. Science 220:261-268, 1983.

Langlotz, C.P. and Shortliffe, E.H.: *Adapting a consultation system to critique user plans*. Int. J. Man-Machine Studies 19:479-496, 1983.

Bischoff, M.B., Shortliffe, E.H., Scott, A.C., Carlson, R.W. and Jacobs, C.D.: *Integration of a computer-based consultant into the clinical setting*. Proceedings 7th Annual Symposium on Computer Applications in Medical Care, pp. 149-152, Baltimore, Maryland, October 1983.

Hickam, D.H., Shortliffe, E.H., Bischoff, M.B., Scott, A.C., Jacobs, C.D. A study of the treatment advice of a computer-based cancer chemotherapy protocol advisor (Memo KSL-85-21). Annals of Internal Medicine 103(6 pt 1):928-936 (1985).

Kent, D.L., Shortliffe, E.H., Carlson, R.W., Bischoff, M.B., Jacobs, C.D. Improvements in data collection through physician use of a computer-based chemotherapy treatment consultant (Memo KSL-85-22). Journal of Clinical Oncology 3:1409-1417 (1985).

Tsuji, S. and Shortliffe, E.H. Graphical access to a medical expert system: I. Design of a knowledge engineer's interface (Memo KSL-85-11). Meth. Inf. Med., Vol. 25, April 1986.

Lane, C.D., Differding, J.C., Shortliffe, E.H. Graphical access to a medical expert system: II. Design of an interface for physicians (Memo KSL-85-15). To appear in Meth. Inf. Med., July 1986.

Langlotz, C., Fagan, L., Tu, S., Williams, J., Sikic, B. ONYX: An architecture for planning in uncertain environments (Memo KSL-85-10). Proceedings of the Ninth International Joint Conference on Artificial Intelligence, pp.447-449, Los Angeles, CA, August 1985.

Stanford Project: PROTEAN Project

Principal Investigators: Oleg Jardetzky
                             (JARDETZKY@SUMEX-AIM.STANFORD.EDU)
                             Nuclear Magnetic Resonance Lab, School of Medicine
                             Stanford University Medical Center
                             Stanford, California 94305

                             Bruce G. Buchanan, Ph.D.
                             (BUCHANAN@SUMEX-AIM.STANFORD.EDU)
                             Computer Science Department
                             Stanford University
                             Stanford, California 94305

Contact Person: Bruce G. Buchanan

The goals of this project are related both to biochemistry and artificial intelligence: (a) use existing AI methods to aid in the determination of the 3-dimensional structure of proteins in solution (not from x-ray crystallography proteins), and (b) use protein structure determination as a test problem for experiments with the AI problem solving structure known as the Blackboard Model. Empirical data from nuclear magnetic resonance (NMR) and other sources may provide enough constraints on structural descriptions to allow protein chemists to bypass the laborious methods of crystallizing a protein and using X-ray crystallography to determine its structure. This problem exhibits considerable complexity; yet there is reason to believe that AI programs can be written that reason much as experts do to resolve these difficulties. A prototype knowledge-based system has been developed that assembles major secondary structures of a protein into families of structures compatible with a given set of distance constraints.

REFERENCES

1. Brinkley, J., Cornelius, C., Altman, R., Hayes-Roth, B. Lichtarge, O., Duncan, B., Buchanan, B.G., Jardetzky, O.: *Application of Constraint Satisfaction Techniques to the Determination of Protein Tertiary Structure.* Report KSL-86-28

2. Buchanan, B.G., Hayes-Roth, B., Lichtarge, O., Altman, A., Brinkley, J., Hewett, M., Cornelius, C., Duncan, B., Jardetzky, O.:*The Heuristic Refinement Method for Deriving Solution Structures of Proteins.* Report KSL-85-41. October 1985. In review for Proceedings of the National Academy of Sciences.

3. Hayes-Roth, B.: *The Blackboard Architecture: A General Framework for Problem Solving?* Report HPP-83-30, Department of Computer Science, Stanford University, 1983.

4. Hayes-Roth, B.: *BB1: An Environment for Building Blackboard Systems that Control, Explain, and Learn about their own Behavior.* Report HPP-84-16, Department of Computer Science, Stanford University, 1984.

5. Hayes-Roth, B.: *A Blackboard Architecture for Control.* Artificial Intelligence 26:251-321, 1985.

6. Hayes-Roth, B. and Hewett, M.: *Learning Control Heuristics in BB1*. Report HPP-85-2, Department of Computer Science, 1985.

7. Jardetzky, O.: *A Method for the Definition of the Solution Structure of Proteins from NMR and Other Physical Measurements: The LAC-Repressor Headpiece*. Proceedings of the International Conference on the Frontiers of Biochemistry and Molecular Biology, Alma Alta, June 17-24, 1984, October, 1984.

8. Hayes-Roth, B., Buchanan, B.G., Lichtarge, O., Hewett, M., Altman, R., Brinkley, J., Cornelius, C., Duncan, B., and Jardetzky, O.: *PROTEAN: Deriving protein structure from constraints*. To appear in Proceedings of the AAAI, 1986.

E. H. Shortliffe

Stanford Project: RADIX -- DERIVING KNOWLEDGE FROM
                              TIME-ORIENTED CLINICAL DATABASES

Principal Investigators: Robert L. Blum, M.D.
Departments of Medicine
and Computer Science
Stanford University
Stanford, California 94305
(415) 497-9421 (BLUM@SUMEX-AIM)

Gio C.M. Wiederhold, Ph.D.
Department of Computer Science
Stanford University
Stanford, California 94305
(415) 497-0685 (WIEDERHOLD@SUMEX-AIM)

The objective of clinical database (DB) systems is to derive medical knowledge from the stored patient observations. However, the process of reliably deriving causal relationships has proven to be quite difficult because of the complexity of disease states and time relationships, strong sources of bias, and problems of missing and outlying data.

The first goal of the RADIX Project is to explore the usefulness of knowledge-based computational techniques in solving this problem of accurate knowledge inference from non-randomized, non-protocol patient records. Central to RADIX is a knowledge base (KB) of medicine and statistics, organized as a taxonomic tree consisting of frames with attached data and procedures. The KB is used to retrieve time-intervals of interest from the DB and to assist with the statistical analysis. Derived knowledge is incorporated automatically into the KB. The American Rheumatism Association DB containing records of 1700 patients is used.

The second goal of the project is to develop a program and set of techniques for automated summarization of patient records. The summarization program is designed to automatically create patient summaries of arbitrary and appropriate complexity as an aid for tasks such as clinical decision making, real-time patient monitoring, surveillance of quality of care, and eventually automated discovery. A prototype summarization module has been implemented in KEE on the Xerox 1108 workstation.

SOFTWARE AVAILABLE ON SUMEX

RADIX--(excluding the knowledge base and clinical database) consists of approximately 400 INTERLISP functions. The following groups of functions may be of interest apart from the RADIX environment:

SPSS Interface Package -- Functions which create SPSS source decks and read SPSS listings from within INTERLISP.

Statistical Tests in INTERLISP -- Translations of the Piezer-Pratt approximations for the T, F, and Chi-square tests into LISP.

Time-Oriented Data Base and Graphics Package -- Autonomous package for maintaining a time-oriented database and displaying labelled time-intervals.

REFERENCES

Monograph

Blum, R.L.: *Discovery and representation of causal relationships from a large time-oriented clinical database: The RX project.* in D.A.B. Lindberg and P.L. Reichertz (Eds.), LECTURE NOTES IN MEDICAL INFORMATICS, Vol. 19, Springer-Verlag, New York, 1982.

Journal Articles

Blum, R.L.: *Computer-Assisted Design of Studies using Routine Clinical Data: Analyzing the Association of Prednisone and Cholesterol.* (Accepted for publication in the Annals of Internal Medicine.)

Blum, R.L.: *Discovery, confirmation, and incorporation of causal relationships from a large time-oriented clinical database: The RX Project.* Computers and Biomedical Research 15(2):164-187, April, 1982.

Conference Proceedings

Downs, S., Walker, M.G. and Blum, R.L.: *Automated Summaries of Patient Medical Records.* Accepted for Medinfo '86.

Walker, M.G., and Blum, R.L.: *Towards Automated Discovery from Clinical Databases: The RADIX Project.* Accepted for Medinfo '86.

National AIM Project: CADUCEUS
                       (INTERNIST-I)

Principal Investigators: Jack D. Myers, M.D. (MYERS@SUMEX-AIM)
                          Harry E. Pople, Ph.D. (POPLE@SUMEX-AIM)
                          University of Pittsburgh
                          Pittsburgh, Pennsylvania 15261
                          Dr. Pople: (412) 624-3490
                          Dr. Myers: (412) 648-9933

The major goal of the INTERNIST-I Project is to produce a reliable and adequately complete diagnostic consultative program in the field of internal medicine. Although this program is intended primarily to aid skilled internists in complicated medical problems, the program may have spin-off as a diagnostic and triage aid to physicians' assistants, rural health clinics, military medicine and space travel. In the design of INTERNIST-I, we have attempted to model the creative, problem-formulation aspect of the clinical reasoning process. The program employs a novel heuristic procedure that composes differential diagnoses, dynamically, on the basis of clinical evidence. During the course of a INTERNIST-1 consultation, it is not uncommon for a number of such conjectured problem foci to be proposed and investigated, with occasional major shifts taking place in the program's conceptualization of the task at hand.

SOFTWARE AVAILABLE ON SUMEX

Versions of INTERNIST-I are available for experimental use, but the project continues to be oriented primarily towards research and development; hence, a stable production version of the system is not yet available for general use.

National AIM Project: CLIPR -- HIERARCHICAL MODELS
                          OF HUMAN COGNITION

Principal Investigators: Walter Kintsch, Ph.D. (KINTSCH@SUMEX-AIM)
                          Peter G. Polson, Ph.D. (POLSON@SUMEX-AIM)
                          Computer Laboratory for Instruction
                          in Psychological Research (CLIPR)
                          Campus Box 345
                          Department of Psychology
                          University of Colorado
                          Boulder, Colorado 80309
                          (303) 492-6991
                          Contact: Dr. Peter G. Polson (Polson@SUMEX-AIM)

The CLIPR Project is concerned with the modeling of complex psychological processes. It is comprised of two research groups. The prose comprehension group has completed a project that carries out the text analysis described by van Dijk & Kintsch (1983), yielding predictions of the recall and readability of that text by human subjects. The human-computer interaction group is developing a quantitative theory of that predicts learning, transfer, and performance for a wide range of computer-tasks, e.g. text editing, Kieras & Polson (1985).

SOFTWARE AVAILABLE ON SUMEX

A set of programs has been developed to perform the microstructure text analysis described in van Dijk & Kintsch (1983) and Kintsch & Greeno (1985). The program accepts a propositionalized text as input, and produces indices that can be used to estimate the text's recall and readability.

REFERENCES

Fletcher, R.C.: Understanding and solving word arithmetic problems: A computer simulation. Technical Report No. 135, Institute of Cognitive Science, Colorado, 1984.

Kieras, D.E. and Polson, P.G.: *The formal analysis of user complexity.* Int. J. Man-Machine Studies, 22, 365-394, 1985.

Kintsch, W. and van Dijk, T.A.: *Toward a model of text comprehension and production.* Psychological Rev. 85:363-394, 1978.

Kintsch, W. and Greeno, J.G.: *Understanding and solving word arithmetic problems.* Psychological Review, 1985, 92, 109-129.

Polson, P.G. and Kieras, D.E.: *A formal description of users' knowledge of how to operate a device and user complexity.* Behavior Research Methods, Instrumentation, & Computers, 1984, 16, 249-255.

Polson, P.G. and Kieras, D.E. *A quantitative model of the learning and performance of text editing knowledge.* In Borman, L. and Curtis, B. (Eds.) Proceedings of the CHI 1985 Conference on Human Factors in Computing. New York: Association for Computing

Machinery. pp. 207-212, 1985.

Polson, P.G. and Jeffries, R. *Instruction in general problem solving skills: An analysis of four approaches.* In (Eds.) Siegel, J., Chipman, S., and Glaser, R. Thinking and learning skills: Relating instructions to basic research: Vol.1. Hillsdale, N.J.: Lawrence Erlbaum Associates, pp. 414-455.

Polson, P.G., Muncher, E., and Engelbeck, G. *Test of a common elements theory of transfer* In Mantei, M. and Orbeton, P. (Eds.) Proceedings of the CHI 1986 Conference on Human Factors in Computing. New York: Association for Computing Machinery. pp. 78-83, 1986.

van Dijk, T.A. and Kintsch, W.: *STRATEGIES OF DISCOURSE COMPREHENSION.* Academic Press, New York, 1983.

Young, S.: A theory and simulation of macrostructure. Technical Report No. 134, Institute of Cognitive Science, Colorado, 1984.

Walker, H.W. & Kintsch, W.: Automatic and strategic aspects of knowledge retrieval. Cognitive Science, 1985, 9, 261-283.

National AIM Project: MENTOR -- MEDICAL EVALUATION OF
                             THERAPEUTIC ORDERS

Principal Investigators: Stuart Speedie, Ph.D.   (SPEEDIE@SUMEX-AIM)
                             School of Pharmacy
                             University of Maryland
                             20 N. Pine Street
                             Baltimore, Maryland   21201
                             (301) 528-7650

                             Terrence F. Blaschke, M.D.   (BLASCHKE@SUMEX-AIM)
                             Department of Medicine
                             Division of Clinical Pharmacology
                             Stanford University Medical Center
                             Stanford, California   94305

Contact: either PI

The goal of the MENTOR project is to implement and begin evaluation of a computer-based methodology for reducing therapeutic misadventures. The project will use principles of artificial intelligence to create an on-line expert system to continuously monitor the drug therapy of individual patients and generate specific warnings of potential and/or actual unintended effects of therapy. The appropriate patient information will be automatically acquired through interfaces to a hospital information system. This data will be monitored by a system that is capable of employing complex chains of reasoning to evaluate therapeutic decisions and arrive at valid conclusions in the context of all information available on the patient. The results reached by the system will be fed back to the responsible physicians to assist future decision making.

Specific objectives of this proposal include:

1. Implement a prototype computer-based expert system to continuously monitor in-patient drug therapy. It will use a modular medical knowledge base and a separate inference engine to apply the knowledge to specific situations.

2. Select a small number of important and frequently occurring drug therapy problems that can lead to therapeutic misadventures and construct a comprehensive knowledge base necessary to detect these situations.

3. Design and begin implementation of an evaluation of the prototype MENTOR system with respect to its impact on the on the physicians' therapeutic decision making as well as its effects on the patient in terms of specific mortality and morbidity measures.

The work in the proposed project will build on the extensive previous work in drug monitoring done by these investigators in the Division of Clinical Pharmacology at Stanford and the University of Maryland School of Pharmacy.

Rutgers AIM Project: RUTGERS RESEARCH RESOURCE-
                                COMPUTERS IN BIOMEDICINE

Principal Investigators: Casimir Kulikowski, Ph.D.
                                Sholom M. Weiss, Ph.D
                                Department of Computer Science
                                Rutgers University
                                New Brunswick, New Jersey 08903
                                (201) 932-2006 (KULIKOWSKI@RUTGERS)
                                (201) 932-2379 (WEISS@RUTGERS)

The Rutgers Research Resource provides research support with artificial intelligence systems, and computing support with its DEC2060 facility to a large number of biomedical scientists and researchers. Research activities are concentrated in two major areas: expert medical systems, models for planning and knowledge acquisition, and general AI systems development.

One of the most significant achievements in bringing the work of the Resource to bear on clinical research and practice lies in the transfer of technology from large machines to microprocessor compatible representations. The initial breakthrough came with the automatic translation of a serum protein electrophoresis interpretation model so that a version could be incorporated in an instrument - a scanning densitometer. It is now being used at several hundred clinical locations.

During the current period, we have been working on a new project with long term implications for the impact of AIM technology: the development of a hand-held microcomputer version of an expert consultation system for front-line health workers. In collaboration with Dr. Chandler Dawson (UCSF), Director of the World Health Organization's Collaborative Centre for the Prevention of Blindness and Trachoma, we have developed a prototype model for consultation on primary eye care. This has been oriented at problems of injury, infection, malnutrition and cataract in situations where an ophthalmologist is unavailable. In most developing nations, the incidence of blindness is 10% to 40% higher than in the USA because of these kinds of problems. With the help of a grant from the USAID, we are developing the systems needed for management of eye disease by front-line health workers in developing nations, and outlying parts of the USA.

REFERENCES

Weiss, S.M. and Kulikowski, C.A. *A Practical Guide to Designing Expert Systems*, Rowman and Allanheld, 1984.

Kulikowski, C.A.*Expert Medical Consultation Systems*, Journal of Medical Systems, v.7, pp. 229-234, 1983.

Weiss, S.M., Kulikowski, C.A., and Galen, R.S.,*Representing Expertise in a Computer Program: The Serum Protein Diagnostic Program*, Journal of Clinical Laboratory Automation, v.3, pp. 383-387, 1983.

Kastner, J., Dawson, C., Weiss, S., Kern, K., Kulikowski, C., *An Expert Consultation System for Frontline Health Workers in Primary Eye Care*, Journal of Medical Systems, Vol. 8, No. 5 (1984).

National AIM Project: SOLVER -- PROBLEM SOLVING EXPERTISE

Principal Investigators: Paul E. Johnson, Ph.D., School of
                               Management and
                               Center for Research in Human Learning
                               205 Elliott Hall
                               University of Minnesota
                               Minneapolis, Minnesota 55455
                               (612) 376-2530 (PJOHNSON@SUMEX-AIM)

                               William B. Thompson, Ph.D.
                               Department of Computer Science
                               136 Lind Hall
                               University of Minnesota
                               Minneapolis, Minnesota 55455
                               (612) 373-0132 (THOMPSON@SUMEX-AIM)

The Minnesota SOLVER project focuses upon the development of strategies for discovering and representing the knowledge and skill of expert problem solvers. Although in the last 15 years considerable progress has been made in synthesizing the expertise required for solving complex problems, most expert systems embody only a limited amount of expertise. What is still lacking is a theoretical framework capable of reducing dependence upon the expert's intuition or on the near exhaustive testing of possible organizations.   Our methodology consists of: (1) extensive use of verbal thinking aloud protocols as a source of information from which to make inferences about underlying knowledge structures and processes; (2) development of computer models as a means of testing the adequacy of inferences derived from protocol studies; (3) testing and refinement of the cognitive models based upon the study of human and model performance in experimental settings.   Currently, we are investigating problem-solving expertise in domains of medicine, financial auditing, management, and law.

SOFTWARE AVAILABLE ON SUMEX

A redesigned version of the Diagnoser simulation model, named Galen, has been implemented on SUMEX.

REFERENCES

Johnson, P.E., Moen, J.B., and Thompson, W.B.: *Garden Path Errors in Medical Diagnosis.* In Bloc, L. and Coombs, M.J. (Eds.), COMPUTER EXPERT SYSTEMS, Springer-Verlag (in press).

Johnson, P.E., Johnson, M.G., and Little, R.K.: *Expertise in trial advocacy: Some considerations for inquiry into its nature and development,* Campbell Law Review, (in press).

Johnson, P.E., "The Expert Mind: A New Challenge for the Information scientist," in *Beyond Productivity: Information System Development for Organizational Effectiveness,* Th. M. A. Bemelmans (editor), Elsevier Science Publishers B.V. (North-Holland), 1984.

Johnson, P.E.: *What kind of expert should a system be?* Journal of Medicine and Philosophy, 8:77-97, 1983.

Johnson, P.E., Duran, A., Hassebrock, F., Moller, J., Prietula, M., Feltovich, P. and Swanson, D.: *Expertise and error in diagnostic*

reasoning. Cognitive Science 5:235-283, 1981.

Thompson, W.B., Johnson, P.E. and Moen, J.B.: *Recognition-based diagnostic reasoning.* Proc. Eighth IJCAI, Karlsruhe, West Germany, August, 1983.

Stanford Project: REFEREE Project

Principal Investigators: Bruce G. Buchanan, Principal Investigator
                             Computer Science Department
                             Stanford University
                             Stanford, California 94305

                             Byron W. Brown, Co-Principal Investigator
                             Department of Medicine
                             Stanford University Medical Center
                             Stanford, California 94305

                             Daniel E. Feldman, Associate Investigator
                             Department of Medicine
                             Stanford University Medical Center
                             Stanford, California 94305

The goals of this project are related both to medical science and artificial intelligence: (a) use AI methods to allow the informed but non-expert reader of the medical literature to evaluate a randomized clinical trial, and (b) use the interpretation of the medical literature as a test problem for studies of knowledge acquisition and fusion of information from disparate sources. REFEREE will be used to evaluate the medical literature of clinical trials to determine the quality of a clinical trial, make judgements on the efficacy of the treatment proposed, and synthesize rules of clinical practice. The research is an initial step toward a more general goal – building computer systems to help the clinician and medical scientist read the medical literature more critically and more rapidly.

E. H. Shortliffe

National AIM Project: Computer-Aided Diagnosis of
                              Malignant Lymph Node Diseases (PATHFINDER)

Principal Investigator: Bharat Nathwani, M.D.
                              Department of Pathology
                              HMR 204
                              2025 Zonal Avenue
                              University of Southern California
                              School of Medicine
                              Los Angeles, California 90033
                              (213) 226-7064 (NATHWANI@SUMEX-AIM)

                              Lawrence M. Fagan, M.D., Ph.D.
                              Department of Medicine
                              Stanford University Medical Center - Room TC135
                              Stanford, California 94305
                              (415) 497-6979 (FAGAN@SUMEX-AIM)

We are building a computer program, called PATHFINDER, to assist in the diagnosis
of lymph node pathology. The project is based at the University of Southern California
in collaboration with the Stanford University Medical Computer Science Group.   A
pilot version of the program provides diagnostic advice on eighty common benign and
malignant diseases of the lymph nodes based on 150 histologic features.   Our research
plans are to develop a full-scale version of the computer program by substantially
increasing the quantity and quality of knowledge and to develop techniques for
knowledge representation and manipulation appropriate to this application area. The
design of the program has been strongly influenced by the INTERNIST/CADUCEUS
program developed on the SUMEX resource.

SOFTWARE AVAILABLE ON SUMEX

PATHFINDER-- A   version   of   the   PATHFINDER   program   is   available   for
             experimentation on the DEC 2060 computer. This version is a pilot
             version of the program, and therefore has not been completely tested.

References

1. Barr, Avron, Paul R. Cohen and Edward A. Feigenbaum. *The Handbook of Artificial Intelligence, Volumes I, II, and III*. William Kaufmann, Inc., Los Altos, CA, 1981 and 1982.

2. Coulter, C. L. "Research Instrument Sharing." *Science 201*, 4354 (1978).

3. Hayes-Roth, F., Waterman, D. A. & Lenat, D. B. (Eds.). *Building Expert Systems*. Addison-Wesley, Reading, MA, 1983.

4. Keith A. Lantz et. al. *V - System 4.1 Reference Manual*. December 1, 1983. Computer Systems Laboratory, Stanford University.

5. Lederberg, J. "Digital Communications and the Conduct of Science: The New Literacy." *Proceedings of the IEEE 66*, 11 (1978).

6. Nilsson, N.. *Principles of Artificial Intelligence*. Tioga, Palo Alto, CA, 1980.

7. Rich, E.. *Artificial Intelligence*. McGraw-Hill, New York, 1983.

8. Winston, P.. *Artificial Intelligence, 2nd ed*. Addison-Wesley, Reading, MA, 1984.

E. H. Shortliffe